(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 July 2002 (25.07.2002)

PCT

(10) International Publication Number
WO 02/057414 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US01/47856

(22) International Filing Date: 22 October 2001 (22.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/241,994 20 October 2000 (20.10.2000) US
60/296,764 8 June 2001 (08.06.2001) US

(71) Applicant *(for all designated States except US)*: **BIOCARDIA, INC.** [US/US]; 384 Oyster Point Boulevard, #4, South San Francisco, CA 94080 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **WOHLGEMUTH, Jay** [US/US]; 664 Hamilton Avenue, Palo Alto, CA 94301 (US). **FRY, Kirk** [US/US]; 2604 Ross Road, Palo Alto, CA 94303 (US). **MATCUK, George** [US/US]; 141C Escondido Village, Stanford, CA 94305 (US). **ALTMAN, Peter** [US/US]; 717 Evelyn Avenue, Albany, CA 94706 (US). **PRENTICE, James** [US/US]; 120 Dolores Street, San Francisco, CA 94103 (US). **PHILLIPS, Julie** [US/US]; 1090 Mirador Terrace, Pacifica, CA 94044 (US). **LY, Ngoc** [US/US]; 2000 Crystal Springs Road 15-14, San Bruno, CA 94066 (US). **WOODWARD, Robert** [US/US]; 1828 Rheem Court, Pleasanton, CA 94588 (US). **QUERTERMOUS, Thomas** [US/US]; 44 El Rey Road, Portola Valley, CA 94028 (US). **JOHNSON, Frances** [US/US]; 44 El Rey Road, Portola Valley, CA 94028 (US).

(74) Agents: **WARD, Michael, R.** et al.; Morrison & Foerster LLP, 425 Market Street, San Francisco, CA 94105-2482 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/057414 A2

(54) Title: LEUKOCYTE EXPRESSION PROFILING

(57) Abstract: Leukocyte gene expression profiling is utilized to identify oligonucleotides from gene expression candidate libraries. The expression libraries are generally immobilized on an array. Diagnostic oligonucleotide sets for analysis of leukocyte-related diseases are described.

LEUKOCYTE EXPRESSION PROFILING

Field of the Invention

This invention is in the field of expression profiling. In particular, this invention is in the field of leukocyte expression profiling.

Background of the Invention

Many of the current shortcomings in diagnosis, prognosis, risk stratification and treatment of disease can be approached through the identification of the molecular mechanisms underlying a disease and through the discovery of nucleotide sequences (or sets of nucleotide sequences) whose expression patterns predict the occurrence or progression of disease states, or predict a patient's response to a particular therapeutic intervention. In particular, identification of nucleotide sequences and sets of nucleotide sequences with such predictive value from cells and tissues that are readily accessible would be extremely valuable. For example, peripheral blood is attainable from all patients and can easily be obtained at multiple time points at low cost. This is a desirable contrast to most other cell and tissue types, which are less readily accessible, or accessible only through invasive and aversive procedures. In addition, the various cell types present in circulating blood are ideal for expression profiling experiments as the many cell types in the blood specimen can be easily separated if desired prior to analysis of gene expression. While blood provides a very attractive substrate for the study of diseases using expression profiling techniques, and for the development of diagnostic technologies and the identification of therapeutic targets, the value of expression profiling in blood samples rests on the degree to which changes in gene expression in these cell types are associated with a predisposition to, and pathogenesis and progression of a disease.

There is an extensive literature supporting the role of leukocytes, e.g., T-and B-lymphocytes, monocytes and granulocytes, including neutrophils, in a wide range of disease processes, including such broad classes as cardiovascular diseases, inflammatory, autoimmune and rheumatic diseases, infectious diseases, transplant rejection, cancer and malignancy, and endocrine diseases. For example, among cardiovascular diseases, such commonly occurring diseases as atherosclerosis, restenosis, transplant vasculopathy and acute coronary syndromes all demonstrate significant T cell involvement (Smith-Norowitz et al. (1999) Clin Immunol 93:168-175; Jude et al. (1994) Circulation 90:1662-8; Belch et al. (1997) Circulation

95:2027-31). These diseases are now recognized as manifestations of chronic inflammatory disorders resulting from an ongoing response to an injury process in the arterial tree (Ross et al. (1999) Ann Thorac Surg 67:1428-33). Differential expression of lymphocyte, monocyte and neutrophil genes and their products has been demonstrated clearly in the literature. Particularly interesting are examples of differential expression in circulating cells of the immune system that demonstrate specificity for a particular disease, such as arteriosclerosis, as opposed to a generalized association with other inflammatory diseases, or for example, with unstable angina rather than quiescent coronary disease.

A number of individual genes, e.g., CD11b/CD18 (Kassirer et al. (1999) Am Heart J 138:555-9); leukocyte elastase (Amaro et al. (1995) Eur Heart J 16:615-22; and CD40L (Aukrust et al. (1999) Circulation 100:614-20) demonstrate some degree of sensitivity and specificity as markers of various vascular diseases. In addition, the identification of differentially expressed target and fingerprint genes isolated from purified populations of monocytes manipulated in various in vitro paradigms has been proposed for the diagnosis and monitoring of a range of cardiovascular diseases, see, e.g., US Patents Numbers 6,048,709; 6,087,477; 6,099,823; and 6,124,433 "COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE" to Falb (*see also*, WO 97/30065). Lockhart, in US Patent Number 6,033,860 "EXPRESSION PROFILES IN ADULT AND FETAL ORGANS" proposes the use of expression profiles for a subset of identified genes in the identification of tissue samples, and the monitoring of drug effects.

The accuracy of technologies based on expression profiling for the diagnosis, prognosis, and monitoring of disease would be dramatically increased if numerous differentially expressed nucleotide sequences, each with a measure of specificity for a disease in question, could be identified and assayed in a concerted manner. In order to achieve this improved accuracy, the appropriate sets of nucleotide sequences need to be identified and validated against numerous samples in combination with relevant clinical data. The present invention addresses these and other needs, and applies to any disease or disease state for which differential regulation of genes, or other nucleotide sequences, of peripheral blood can be demonstrated.

Summary of the Invention

The present invention is thus directed to a system for detecting differential gene expression. In one format, the system has one or more isolated DNA molecules

wherein each isolated DNA molecule detects expression of a gene selected from the group of genes corresponding to the oligonucleotides depicted in the Sequence Listing. It is understood that the DNA sequences and oligonucleotides of the invention may have slightly different sequences that those identified herein. Such sequence variations are understood to those of ordinary skill in the art to be variations in the sequence which do not significantly affect the ability of the sequences to detect gene expression.

The sequences encompassed by the invention have at least 40-50, 50-60, 70-80, 80-85, 85-90, 90-95 % or 95-100% sequence identity to the sequences disclosed herein. In some embodiments, DNA molecules are less than about any of the following lengths (in bases or base pairs): 10,000; 5,000; 2500; 2000; 1500; 1250; 1000; 750; 500; 300; 250; 200; 175; 150; 125; 100; 75; 50; 25; 10. In some embodiments, DNA molecule is greater than about any of the following lengths (in bases or base pairs): 10; 15; 20; 25; 30; 40; 50; 60; 75; 100; 125; 150; 175; 200; 250; 300; 350; 400; 500; 750; 1000; 2000; 5000; 7500; 10000; 20000; 50000. Alternately, a DNA molecule can be any of a range of sizes having an upper limit of 10,000; 5,000; 2500; 2000; 1500; 1250; 1000; 750; 500; 300; 250; 200; 175; 150; 125; 100; 75; 50; 25; or 10 and an independently selected lower limit of 10; 15; 20; 25; 30; 40; 50; 60; 75; 100; 125; 150; 175; 200; 250; 300; 350; 400; 500; 750; 1000; 2000; 5000; 7500 wherein the lower limit is less than the upper limit.

The gene expression system may be a candidate library, a diagnostic agent, a diagnostic oligonucleotide set or a diagnostic probe set. The DNA molecules may be genomic DNA, protein nucleic acid (PNA), cDNA or synthetic oligonucleotides.

In one format, the gene expression system is immobilized on an array. The array may be a chip array, a plate array, a bead array, a pin array, a membrane array, a solid surface array, a liquid array, an oligonucleotide array, a polynucleotide array, a cDNA array, a microfilter plate, a membrane or a chip.

In one format, the genes detected by the gene expression system are selected from the group of genes corresponding to the oligonucleotides depicted in SEQ ID NO:2476, SEQ ID NO: 2407, SEQ ID NO:2192, SEQ ID NO: 2283, SEQ ID NO:6025, SEQ ID NO: 4481, SEQ ID NO:3761, SEQ ID NO: 3791, SEQ ID NO:4476, SEQ ID NO: 4398, SEQ ID NO:7401, SEQ ID NO: 1796, SEQ ID NO:4423, SEQ ID NO: 4429, SEQ ID NO:4430, SEQ ID NO: 4767, SEQ ID NO:4829 and SEQ ID NO: 8091.

The present invention is further directed to a diagnostic agent comprising an oligonucleotide wherein the oligonucleotide has a nucleotide sequence selected from the Sequence Listing wherein the oligonucleotide detects expression of a gene that is differentially expressed in leukocytes in an individual over time. In one format, the oligonucleotide has a nucleotide sequence selected from the group consisting of SEQ ID NO:2476, SEQ ID NO: 2407, SEQ ID NO:2192, SEQ ID NO: 2283, SEQ ID NO:6025, SEQ ID NO: 4481, SEQ ID NO:3761, SEQ ID NO: 3791, SEQ ID NO:4476, SEQ ID NO: 4398, SEQ ID NO:7401, SEQ ID NO: 1796, SEQ ID NO:4423, SEQ ID NO: 4429, SEQ ID NO:4430, SEQ ID NO: 4767, SEQ ID NO:4829 and SEQ ID NO: 8091

The present invention is futher directed to a system for detecting gene expression in leukocytes comprising an isolated DNA molecule wherein the isolated DNA molecule detects expression of a gene wherein the gene is selected from the group of genes corresponding to the oligonucleotides depicted in the Sequence Listing and the gene is differentially expressed in the leukocytes in an individual with at least one disease criterion for a disease selected from Table 1 as compared to the expression of the gene in leukocytes in an individual without the at least one disease criterion.

The present invention is further directed to a gene expression candidate library comprising at least two oligonucleotides wherein the oligonucleotides have a sequence selected from those oligonucleotide sequences listed in Table 2, Table 3, and the Sequence Listing. Table 3 encompasses Tables 3A, 3B and 3C. The oligonucleotides of the candidate library may comprise deoxyribonucleic acid (DNA), ribonucleic acid (RNA), protein nucleic acid (PNA), synthetic oligonucleotides, or genomic DNA.

In one embodiment, the candidate library is immobilized on an array. The array may comprises one or more of: a chip array, a plate array, a bead array, a pin array, a membrane array, a solid surface array, a liquid array, an oligonucleotide array, a polynucleotide array or a cDNA array, a microtiter plate, a pin array, a bead array, a membrane or a chip. Individual members of the libraries are may be separately immobilized.

The present invention is further directed to a diagnostic oligonucleotide set for a disease having at least two oligonucleotides wherein the oligonucleotides have a sequence selected from those oligonucleotide sequences listed in Table 2, Table 3, or

the Sequence Listing which are differentially expressed in leukocytes genes in an individual with at least one disease criterion for at least one leukocyte-related disease as compared to the expression in leukocytes in an individual without the at least one disease criterion, wherein expression of the two or more genes of the gene expression library is correlated with at least one disease criterion.

The present invention is further directed to a diagnostic oligonucleotide set for a disease having at least one oligonucleotide wherein the oligonucleotide has a sequence selected from those sequences listed in Table 2, Table 3, or the sequence listing which is differentially expressed in leukocytes in an individual with at least one disease criterion for a disease selected from Table 1 as compared toleukocytes in an individual without at least one disease criterion, wherein expression of the at least one gene from the gene expression library is correlated with at least one disease criterion, wherein the differential expression of the at least one gene has not previously been described. In one format, two or more oligonucleotides are utilized.

In the diagnostic oligonucleotide sets of the invention the disease criterion may include data selected from patient historic, diagnostic, prognostic, risk prediction, therapeutic progress, and therapeutic outcome data. This includes lab results, radiology results, pathology results such as histology, cytology and the like, physical examination findings, and medication lists.

In the diagnostic oligonucleotide sets of the invention the leukocytes comprise peripheral blood leukocytes or leukocytes derived from a non-blood fluid. The non-blood fluid may be selected from colon, sinus, spinal fluid, saliva, lymph fluid, esophagus, small bowel, pancreatic duct, biliary tree, ureter, vagina, cervix uterus and pulmonary lavage fluid.

In the diagnostic oligonucleotide sets of the invention the leukocytes may include leukocytes derived from urine or a joint biopsy sample or biopsy of any other tissue or may be T-lymphocytes.

In the diagnostic oligonucleotide sets of the invention the disease may be selected from cardiac allograft rejection, kidney allograft rejection, liver allograft rejection, atherosclerosis, congestive heart failure, systemic lupus erythematosis (SLE), rheumatoid arthritis, osteoarthritis, and cytomegalovirus infection.

The diagnostic oligonucleotide sets of the invention may further include one or more cytomegalovirus (CMV) nucleotide sequences, wherein expression of the CMV nucleotide sequence is correlated with CMV infection.

The diagnostic nucleotide sets of the invention may further include one or more Epstein-Barr virus (EBV) nucleotide sequences, wherein expression of the one or more EBV nucleotide sequences is correlated with EBV infection.

In the present invention, expression may be differential expression, wherein the differential expression is one or more of a relative increase in expression, a relative decrease in expression, presence of expression or absence of expression, presence of disease or absence of disease. The differential expression may be RNA expression or protein expression. The differential expression may be between two or more samples from the same patient taken on separate occasions or between two or more separate patients or between two or more genes relative to each other.

The present invention is further directed to a diagnostic probe set for a disease where the probes correspond to at least one oligonucleotide wherein the oligonucleotides have a sequence ssuch as those listed in Table 2, Table 3, or the Sequence Listing which is differentially expressed in leukocytes in an individual with at least one disease criterion for a disease selected from Table 1 as comapared to leukocytes in an individual without the at least one disease criterion, wherein expression of the oligonucleotide is correlated with at least one disease criterion, and further wherein the differential expression of the at least one nucleotide sequence has not previously been described.

The present invention is further directed to a diagnostic probe set wherein the probes include one or more of probes useful for proteomics and probes for nucleic acids cDNA, or synthetic oligonucleotides.

The present invention is further directed to an isolated nucleic acid having a sequences such as those listed in Table 3B or Table 3C or the Sequence Listing.

The present invention is further directed to polypeptides wherein the polypeptides are encoded by the nucleic acid sequences in Tables 3B, 3C and the Sequence Listing.

The present invention is further directed to a polynucleotide expression vector containing the polynucleotide of Tables 3B-3C or the Sequence Listing in operative association with a regulatory element which controls expression of the polynucleotide in a host cell. The present invention is further directed to host cells transformed with the expression vectors of the invention. The host cell may be prokaryotic or eukaryotic.

The present invention is further directed to fusion proteins produced by the host cells of the invention. The present invention is further directed to antibodies directed to the fusion proteins of the invention. The antibodies may be monoclonal or polyclonal antibodies.

The present invention is further directed to kits comprising the diagnostic oligonucleotide sets of the invention. The kits may include instructions for use of the kit.

The present invention is further directed to a method of diagnosing a disease by obtaining a leukocyte sample from an individual, hybridizing nucleic acid derived from the leukocyte sample with a diagnostic oligonucleotide set, and comparing the expression of the diagnostic oligonucleotide set with a molecular signature indicative of the presence or absence of the disease.

The present invention is further directed to a method of detecting gene expression by a) isolating RNA and b) hybridizing the RNA to isolated DNA molecules wherein the isolated DNA molecules detect expression of a gene wherein the gene corresponds to one of the oligonucleotides depicted in the Sequence Listing.

The present invention is further directed to a method of detecting gene expression by a) isolating RNA; b) converting the RNA to nucleic acid derived from the RNA and c) hybridizing the nucleic acid derived from the RNA to isolated DNA molecules wherein the isolated DNA molecules detect expression of a gene wherein the gene corresponds to one of the oligonucleotides depicted in the Sequence Listing. In one format, the nucleic acid derived from the RNA is cDNA.

The present invention is further directed to a method of detecting gene expression by a) isolating RNA; b) converting the RNA to cRNA or aRNA and c) hybridizing the cRNA or aRNA to isolated DNA molecules wherein the isolated DNA molecules detect expression of a gene corresponding to one of the oligonucleotides depicted in the Sequence Listing.

The present invention is further directed to a method of monitoring progression of a disease by obtaining a leukocyte sample from an individual, hybridizing the nucleic acid derived from leukocyte sample with a diagnostic oligonucleotide set, and comparing the expression of the diagnostic oligonucleotide set with a molecular signature indicative of the presence or absence of disease progression.

The present invention is further directed to a method of monitoring the rate of progression of a disease by obtaining a leukocyte sample from an individual, hybridizing the nucleic acid derived from leukocyte sample with a diagnostic oligonucleotide set, and comparing the expression of the diagnostic oligonucleotide set with a molecular signature indicative of the presence or absence of disease progression.

The present invention is further directed to a method of predicting therapeutic outcome by obtaining a leukocyte sample from an individual, hybridizing the nucleic acid derived from leukocyte sample with a diagnostic oligonucleotide set, and comparing the expression of the diagnostic oligonucleotide set with a molecular signature indicative of the predicted therapeutic outcome.

The present invention is further directed to a method of determining prognosis by obtaining a leukocyte sample from an individual, hybridizing the nucleic acid derived from leukocyte sample with a diagnostic oligonucleotide set, and comparing the expression of the diagnostic oligonucleotide set with a molecular signature indicative of the prognosis.

The present invention is further directed to a method of predicting disease complications by obtaining a leukocyte sample from an individual, hybridizing nucleic acid derived from the leukocyte sample with a diagnostic oligonucleotide set, and comparing the expression of the diagnostic oligonucleotide set with a molecular signature indicative of the presence or absence of disease complications.

The present invention is further directed to a method of monitoring response to treatment, by obtaining a leukocyte sample from an individual, hybridizing the nucleic acid derived from leukocyte sample with a diagnostic oligonucleotide set, and comparing the expression of the diagnostic oligonucleotide set with a molecular signature indicative of the presence or absence of response to treatment.

In the methods of the invention the invention may further include characterizing the genotype of the individual, and comparing the genotype of the individual with a diagnostic genotype, wherein the diagnostic genotype is correlated with at least one disease criterion. The genotype may be analyzed by one or more methods selected from the group consisting of Southern analysis, RFLP analysis, PCR, single stranded conformation polymorphism and SNP analysis.

The present invention is further directed to a method of non-invasive imaging by providing an imaging probe for a nucleotide sequence that is differentially

expressed in leukocytes from an individual with at least one disease criterion for at least one leukocyte-implicated disease where leukocytes localize at the site of disease, wherein the expression of the at least one nucleotide sequence is correlated with the at least one disease criterion by (a) contacting the probe with a population of leukocytes; (b) allowing leukocytes to localize to the site of disease or injury and (c) detecting an image.

The present invention is further directed to a control RNA for use in expression profile analysis, where the RNA extracted from the buffy coat samples isfrom at least four individuals.

The present invention is further directed to a method of collecting expression profiles, comprising comparing the expression profile of an individual with the expression profile of buffy coat control RNA, and analyzing the profile.

The present invention is further directed to a method of RNA preparation suitable for diagnostic expression profiling by obtaining a leukocyte sample from a subject, adding actinomycin-D to a final concentration of 1 ug/ml, adding cycloheximide to a final concentration of 10 ug/ml, and extracting RNA from the leukocyte sample. In the method of RNA preparation of the invention the actinomycin-D and cycloheximide may be present in a sample tube to which the leukocyte sample is added. The method may further include centrifuging the sample at 4°C to separate mononuclear cells.

The present invention is further directed to a leukocyte oligonucleotide set including at least two oligonucleotides which are differentially expressed in leukocytes undergoing adhesion to an endothelium relative to expression in leukocytes not undergoing adhesion to an endothelium, wherein expression of the two oligonucleotides is correlated with the at least one indicator of adhesion state.

The present invention is further directed to a method of identifying at least one diagnostic probe set for assessing atherosclerosis by (a) providing a library of candidate oligonucleotides, which candidate oligonucleotides are differentially expressed in leukocytes which are undergoing adhesion to an endothelium relative to their expression in leukocytes that are not undergoing adhesion to an endothelium; (b) assessing expression of two or more oligonucleotides, which two or more oligonucleotides correspond to components of the library of candidate oligonucleotides, in a subject sample of leukocytes; (c) correlating expression of the two or more oligonucleotides with at least one criterion, which criterion includes one

or more indicators of adhesion to an endothelium; and, (d) recording the molecular signature in a database.

The present invention is further directed to a method of identifying at least one diagnostic probe set for assessing atherosclerosis by (a) providing a library of candidate oligonucleotides, which candidate oligonucleotides are differentially expressed in leukocytes which are undergoing adhesion to an endothelium relative to their expression in leukocytes that are not undergoing adhesion to an endothelium; (b) assessing expression of two or more oligonucleotides, which two or more oligonucleotides correspond to components of the library of candidate nucleotide sequences, in a subject sample of epithelial cells; (c) correlating expression of the two or more nucleotide sequences with at least one criterion, which criterion comprises one or more indicator of adhesion to an endothelium; and(d) recording the molecular signature in a database.

The present invention is further directed to methods of leukocyte expression profiling including methods of analyzing longitudinal clinical and expression data. The rate of change and/or magnitude and direction of change of gene expression can be correlated with disease states and the rate of change of clinical conditions/data and/or the magnitude and direction of changes in clinical data. Correlations may be discovered by examining these expression or clinical changes that are not found in the absence of such changes.

The present invention is further directed to methods of leukocyte profiling for analysis and/or detection of one or more viruses. The virus may be CMV, HIV, hepatitis or other viruses. Both viral and human leukocyte genes can be subjected to expression profiling for these purposes.

Brief Description of the Sequence Listing

The table below gives a description of the sequence listing. There are 8830 entries. The Sequence Listing presents 50mer oligonucleotide sequences derived from human leukocyte, plant and viral genes. These are listed as SEQ IDs 1-8143. The 50mer sequences and their sources are also displayed in Table 8. Most of these 50mers were designed from sequences of genes in Tables 2, 3A, B and C and the Sequence listing.

SEQ IDs 8144-8766 are the cDNA sequences derived from human leukocytes that were not homologous to UniGene sequences or sequences found in dbEST at the

time they were searched. Some of these sequences match human genomic sequences and are listed in Tables 3B and C. The remaining clones are putative cDNA sequences that contained less than 50% masked nucleotides when submitted to RepeatMasker, were longer than 147 nucleotides, and did not have significant similarity to the UniGene Unique database, dbEST, the NR nucleotide database of Genbank or the assembled human genome of Genbank.

SEQ IDs 8767-8770, 8828-8830 and 8832are sequences that appear in the text and examples (primer, masked sequences, exemplary sequences, etc.).

SEQ IDs 8771-8827 are CMV PCR primers described in Example 17.

Brief Description of the Figures

Figure 1: Figure 1 is a schematic flow chart illustrating a schematic instruction set for characterization of the nucleotide sequence and/or the predicted protein sequence of novel nucleotide sequences.

Figure 2: Figure 2 depicts the components of an automated RNA preparation machine.

Figure 3: Figure 3 describes kits useful for the practice of the invention. Figure 3A describes the contents of a kit useful for the discovery of diagnostic nucleotide sets. Figure 3B describes the contents of a kit useful for the application of diagnostic nucleotide sets.

Figure 4 shows the results of six hybridizations on a mini array graphed (n=6 for each column). The error bars are the SEM. This experiment shows that the average signal from AP prepared RNA is 47% of the average signal from GS prepared RNA for both Cy3 and Cy5.

Figure 5 shows the average background subtracted signal for each of nine leukocyte-specific genes on a mini array. This average is for 3-6 of the above-described hybridizations for each gene. The error bars are the SEM.

Figure 6 shows the ratio of Cy3 to Cy5 signal for a number of genes. After normalization, this ratio corrects for variability among hybridizations and allows comparison between experiments done at different times. The ratio is calculated as the Cy3 background subtracted signal divided by the Cy5 background subtracted signal. Each bar is the average for 3-6 hybridizations. The error bars are SEM.

Figure 7 shows data median Cy3 background subtracted signals for control RNAs using mini arrays.

Figure 8 shows data from an array hybridization.

Figure 9 shows a comparison of gene expression in samples obtained from cardiac transplant patients wth low rejection grade and high rejection grade.

Figure 10 shows differential gene expression between samples from patients with grade 0 and grade 3A rejection.

Brief Description of the Tables

Table 1: Table 1 lists diseases or conditions amenable to study by leukocyte profiling.

Table 2: Table 2 describes genes and other nucleotide sequences identified using data mining of publically available publication databases and nucleotide sequence databases. Corresponding Unigene (build 133) cluster numbers are listed with each gene or other nucleotide sequence.

Table 3A: Table 3A describes 48 clones whose sequences align to two or more non-contiguous sequences on the same assembled human contig of genomic sequence. The Accession numbers are from the March 15, 2001 build of the human genome. The file date for the downloaded data was 4/17/01. The alignments of the clone and the contig are indicated in the table. The start and stop offset of each matching region is indicated in the table. The sequence of the clones themselves is included in the sequence listing. The alignments of these clones strongly suggest that they are novel nucleotide sequences. Furthermore, no EST or mRNA aligning to the clone was found in the database. These sequences may prove useful for the prediction of clinical outcomes.

Table 3B: Table 3B describes Identified Genomic Regions that code for novel mRNAs. The table contains 591 identified genomic regions that are highly similar to the cDNA clones. Those regions that are within ~100 to 200 Kb of each other on the same contig are likely to represent exons of the same gene. The indicated clone is exemplary of the cDNA clones that match the indicated genomic region. The "number clones" column indicates how many clones were isolated from the libraries that are similar to the indicated region of the chromosome. The probability number is the likelihood that region of similarity would occur by chance on a random sequence. The Accession numbers are from the March 15, 2001 build of the human genome. The file date for the downloaded data was 4/17/01. These sequences may prove useful for the prediction of clinical outcomes.

Table 3C: Table 3C describes differentially expressed nucleotide sequences useful for the prediction of clinical outcomes. This table contains 4517 identified cDNAs and cDNA regions of genes that are members of a leukocyte candidate library, for use in measuring the expression of nucleotide sequences that could subsequently be correlated with human clinical conditions. The regions of similarity were found by searching three different databases for pair wise similarity using blastn. The three databases were UniGene Unique build 3/30/01, file Hs.seq.uniq.Z; the downloadable database at ftp.ncbi.nlm.nih.com/blast/db/est human.Z with date 4/8/01 which is a section of Genbank version 122; and the non-redundant section of Genbank ver 123. The Hs.XXXXX numbers represent UniGene accession numbers from the Hs.seq.uniq.Z file of 3/30/01. The clone sequences are not in the sequence listing.

Table 4: Table 4 describes patient groups and diagnostic gene sets

Table 5: Table 5 describes the nucleotide sequence databases used in the sequence analysis described herein.

Table 6: Table 6 describes the algorithms and software packages used for exon and polypeptide prediction used in the sequence analysis described herein.

Table 7: Table 7 describes the databases and algorithms used for the protein sequence analysis described herein.

Table 8: Table 8 describes leukocyte probes spotted on the microarrays.

Table 9: Table 9 describes Cardiac Transplant patient RNA samples and array hybridizations.

Table 10: Table 10 describes differentially expressed probes identified when comparing leukocyte expression profiles obtained from high and low grade cardiac transplant rejection patients.

Detailed Description of the Invention

Definitions

Unless defined otherwise, all scientific and technical terms are understood to have the same meaning as commonly used in the art to which they pertain. For the purpose of the present invention, the following terms are defined below.

In the context of the invention, the term "gene expression system" refers to any system, device or means to detect gene expression and includes diagnostic agents, candidate libraries, oligonucleotide sets or probe sets.

The term "diagnostic oligonucleotide set" generally refers to a set of two or more oligonucleotides that, when evaluated for differential expression of their products, collectively yields predictive data. Such predictive data typically relates to diagnosis, prognosis, monitoring of therapeutic outcomes, and the like. In general, the components of a diagnostic oligonucleotide set are distinguished from nucleotide sequences that are evaluated by analysis of the DNA to directly determine the genotype of an individual as it correlates with a specified trait or phenotype, such as a disease, in that it is the pattern of expression of the components of the diagnostic nucleotide set, rather than mutation or polymorphism of the DNA sequence that provides predictive value. It will be understood that a particular component (or member) of a diagnostic nucleotide set can, in some cases, also present one or more mutations, or polymorphisms that are amenable to direct genotyping by any of a variety of well known analysis methods, e.g., Southern blotting, RFLP, AFLP, SSCP, SNP, and the like.

A "disease specific target oligonucleotide sequence" is a gene or other oligonucleotide that encodes a polypeptide, most typically a protein, or a subunit of a multi-subunit protein, that is a therapeutic target for a disease, or group of diseases.

A "candidate library" or a "candidate oligonucleotide library" refers to a collection of oligonucleotide sequences (or gene sequences) that by one or more criteria have an increased probability of being associated with a particular disease or group of diseases. The criteria can be, for example, a differential expression pattern in a disease state or in activated or resting leukocytes in vitro as reported in the scientific or technical literature, tissue specific expression as reported in a sequence database, differential expression in a tissue or cell type of interest, or the like. Typically, a candidate library has at least 2 members or components; more typically, the library has in excess of about 10, or about 100, or about 1000, or even more, members or components.

The term "disease criterion" is used herein to designate an indicator of a disease, such as a diagnostic factor, a prognostic factor, a factor indicated by a medical or family history, a genetic factor, or a symptom, as well as an overt or confirmed diagnosis of a disease associated with several indicators such as those selected from the above list. A disease criterian includes data describing a patient's health status, including retrospective or prospective health data, e.g. in the form of the

patient's medical history, laboratory test results, diagnostic test result, clinical events, medications, lists, response(s) to treatment and risk factors, etc.

The terms "molecular signature" or "expression profile" refers to the collection of expression values for a plurality (e.g., at least 2, but frequently about 10, about 100, about 1000, or more) of members of a candidate library. In many cases, the molecular signature represents the expression pattern for all of the nucleotide sequences in a library or array of candidate or diagnostic nucleotide sequences or genes. Alternatively, the molecular signature represents the expression pattern for one or more subsets of the candidate library. The term "oligonucleotide" refers to two or more nucleotides. Nucleotides may be DNA or RNA, naturally occurring or synthetic.

The term "healthy individual," as used herein, is relative to a specified disease or disease criterion. That is, the individual does not exhibit the specified disease criterion or is not diagnosed with the specified disease. It will be understood, that the individual in question, can, of course, exhibit symptoms, or possess various indicator factors for another disease.

Similarly, an "individual diagnosed with a disease" refers to an individual diagnosed with a specified disease (or disease criterion). Such an individual may, or may not, also exhibit a disease criterion associated with, or be diagnosed with another (related or unrelated) disease.

An "array" is a spatially or logically organized collection, e.g., of oligonucleotide sequences or nucleotide sequence products such as RNA or proteins encoded by an oligonucleotide sequence. In some embodiments, an array includes antibodies or other binding reagents specific for products of a candidate library.

When referring to a pattern of expression, a "qualitative" difference in gene expression refers to a difference that is not assigned a relative value. That is, such a difference is designated by an "all or nothing" valuation. Such an all or nothing variation can be, for example, expression above or below a threshold of detection (an on/off pattern of expression). Alternatively, a qualitative difference can refer to expression of different types of expression products, e.g., different alleles (e.g., a mutant or polymorphic allele), variants (including sequence variants as well as post-translationally modified variants), etc.

In contrast, a "quantitative" difference, when referring to a pattern of gene expression, refers to a difference in expression that can be assigned a value on a

graduated scale, (e.g., a 0-5 or 1-10 scale, a + - +++ scale, a grade 1- grade 5 scale, or the like; it will be understood that the numbers selected for illustration are entirely arbitrary and in no-way are meant to be interpreted to limit the invention).

Gene Expression Systems of the Invention

The invention is directed to a gene expression system having one or more oligonucleotides wherein the one or more oligonucleotides has a nucleotide sequence which detects expression of a gene corresponding to the oligonucleotides depicted in the Sequence Listing. In one format, the oligonucleotide detects expression of a gene that is differentially expressed in leukocytes. The gene expression system may be a candidate library, a diagnostic agent, a diagnostic oligonucleotide set or a diagnostic probe set. The DNA molecules may be genomic DNA, protein nucleic acid (PNA), cDNA or synthetic oligonucleotides. Following the procedures taught herein, one can identity sequences of interest for analyzing gene expression in leukocytes. Such sequences may be predictive of a disease state.

Diagnostic oligonucleotides of the invention

The invention relates to diagnostic nucleotide set(s) comprising members of the leukocyte candidate library listed in Table 2, Table 3 and in the Sequence Listing, for which a correlation exists between the health status of an individual, and the individual's expression of RNA or protein products corresponding to the nucleotide sequence. In some instances, only one oligonucleotide is necessary for such detection. Members of a diagnostic oligonucleotide set may be identified by any means capable of detecting expression of RNA or protein products, including but not limited to differential expression screening, PCR, RT-PCR, SAGE analysis, high-throughput sequencing, microarrays, liquid or other arrays, protein-based methods (e.g., western blotting, proteomics, and other methods described herein), and data mining methods, as further described herein.

In one embodiment, a diagnostic oligonucleotide set comprises at least two oligonucleotide sequences listed in Table 2 or Table 3 or the Sequence Listing which are differentially expressed in leukocytes in an individual with at least one disease criterion for at least one leukocyte-implicated disease relative to the expression in individual without the at least one disease criterion, wherein expression of the two or more nucleotide sequences is correlated with at least one disease criterion, as described below. In another embodiment, a diagnostic nucleotide set comprises

at least one oligonucleotide having an oligonucleotide sequence listed in Table 2 or 3 or the Sequence Listing which is differentially expressed, and further wherein the differential expression/correlation has not previously been described. In some embodiments, the diagnostic nucleotide set is immobilized on an array.

The invention also provides diagnostic probe sets. It is understood that a probe includes any reagent capable of specifically identifying a nucleotide sequence of the diagnostic nucleotide set, including but not limited to a DNA, a RNA, cDNA, synthetic oligonucleotide, partial or full-length nucleic acid sequences. In addition, the probe may identify the protein product of a diagnostic nucleotide sequence, including, for example, antibodies and other affinity reagents. It is also understood that each probe can correspond to one gene, or multiple probes can correspond to one gene, or both, or one probe can correspond to more than one gene.

Homologs and variants of the disclosed nucleic acid molecules may be used in the present invention. Homologs and variants of these nucleic acid molecules will possess a relatively high degree of sequence identity when aligned using standard methods. The sequences encompassed by the invention have at least 40-50, 50-60, 70-80, 80-85, 85-90, 90-95 or 95-100% sequence identity to the sequences disclosed herein.

It is understood that for expression profiling, variations in the disclosed sequences will still permit detection of gene expression. The degree of sequence identity required to detect gene expression varies depending on the length of the oligomer. For a 60 mer, 6-8 random mutations or 6-8 random deletions in a 60 mer do not affect gene expression detection. Hughes, TR, et al. "Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nature Biotechnology, 19:343-347(2001). As the length of the DNA sequence is increased, the number of mutations or deletions permitted while still allowing gene expression detection is increased.

As will be appreciated by those skilled in the art, the sequences of the present invention may contain sequencing errors. That is, there may be incorrect nucleotides, frameshifts, unknown nucleotides, or other types of sequencing errors in any of the sequences; however, the correct sequences will fall within the homology and stringency definitions herein.

The minimum length of an oligonucleotide probe necessary for specific hybridization in the human genome can be estimated using two approaches. The first method uses a statistical argument that the probe will be unique in the human genome by chance. Briefly, the number of independent perfect matches (Po) expected for an oligonucleotide of length L in a genome of complexity C can be calculated from the equation (Laird CD, Chromosoma 32:378 (1971):

$Po=(1/4)^L * 2C$

In the case of mammalian genomes, $2C = \sim 3.6 \times 10^9$, and an oligonucleotide of 14-15 nucleotides is expected to be represented only once in the genome. However, the distribution of nucleotides in the coding sequence of mammalian genomes is nonrandom (Lathe, R. J. Mol. Biol. 183:1 (1985) and longer oligonucleotides may be preferred in order to in increase the specificity of hybridization. In practical terms, this works out to probes that are 19-40 nucleotides long (Sambrook J et al., infra). The second method for estimating the length of a specific probe is to use a probe long enough to hybridize under the chosen conditions and use a computer to search for that sequence or close matches to the sequence in the human genome and choose a unique match. Probe sequences are chosen based on the desired hybridization properties as described in Chapter 11 of Sambrook et al, infra. The PRIMER3 program is useful for designing these probes (S. Rozen and H. Skaletsky 1996,1997; Primer3 code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html). The sequences of these probes are then compared pair wise against a database of the human genome sequences using a program such as BLAST or MEGABLAST (Madden, T.L et al.(1996) Meth. Enzymol. 266:131-141). Since most of the human genome is now contained in the database, the number of matches will be determined. Probe sequences are chosen that are unique to the desired target sequence.

In some embodiments, a diagnostic probe set is immobilized on an array. The array is optionally comprises one or more of: a chip array, a plate array, a bead array, a pin array, a membrane array, a solid surface array, a liquid array, an oligonucleotide array, a polynucleotide array or a cDNA array, a microtiter plate, a pin array, a bead array, a membrane or a chip.

In some embodiments, the leukocyte-implicated disease is selected from the diseases listed in Table 1. In other embodiments, the disease is atherosclerosis or

cardiac allograft rejection. In other embodiments, the disease is congestive heart failure, angina, myocardial infarction, systemic lupus erythematosis (SLE) and rheumatoid arthritis.

General Molecular Biology References

In the context of the invention, nucleic acids and/or proteins are manipulated according to well known molecular biology techniques. Detailed protocols for numerous such procedures are described in, e.g., in Ausubel et al. Current Protocols in Molecular Biology (supplemented through 2000) John Wiley & Sons, New York ("Ausubel"); Sambrook et al. Molecular Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook"), and Berger and Kimmel Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA ("Berger").

In addition to the above references, protocols for in vitro amplification techniques, such as the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Q-replicase amplification, and other RNA polymerase mediated techniques (e.g., NASBA), useful e.g., for amplifying cDNA probes of the invention, are found in Mullis et al. (1987) U.S. Patent No. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press Inc. San Diego, CA (1990) ("Innis"); Arnheim and Levinson (1990) C&EN 36; The Journal Of NIH Research (1991) 3:81; Kwoh et al. (1989) Proc Natl Acad Sci USA 86, 1173; Guatelli et al. (1990) Proc Natl Acad Sci USA 87:1874; Lomell et al. (1989) J Clin Chem 35:1826; Landegren et al. (1988) Science 241:1077; Van Brunt (1990) Biotechnology 8:291; Wu and Wallace (1989) Gene 4: 560; Barringer et al. (1990) Gene 89:117, and Sooknanan and Malek (1995) Biotechnology 13:563. Additional methods, useful for cloning nucleic acids in the context of the present invention, include Wallace et al. U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369:684 and the references therein.

Certain polynucleotides of the invention, e.g., oligonucleotides can be synthesized utilizing various solid-phase strategies involving mononucleotide- and/or trinucleotide-based phosphoramidite coupling chemistry. For example, nucleic acid sequences can be synthesized by the sequential addition of activated monomers and/or

trimers to an elongating polynucleotide chain. See e.g., Caruthers, M.H. et al. (1992) Meth Enzymol 211:3.

In lieu of synthesizing the desired sequences, essentially any nucleic acid can be custom ordered from any of a variety of commercial sources, such as The Midland Certified Reagent Company (mcrc@oligos.com), The Great American Gene Company (www.genco.com), ExpressGen, Inc. (www.expressgen.com), Operon Technologies, Inc. (www.operon.com), and many others.

Similarly, commercial sources for nucleic acid and protein microarrays are available, and include, e.g., Agilent Technologies, Palo Alto, CA (http://www.agilent.com/) Affymetrix, Santa Clara,CA (http://www.affymetrix.com/); and Incyte, Palo Alto, CA (http://www.incyte.com/) and others.

Identification of diagnostic nucleotide sets

Candidate library

Libraries of candidates that are differentially expressed in leukocytes are substrates for the identification and evaluation of diagnostic oligonucleotide sets and disease specific target nucleotide sequences.

The term leukocyte is used generically to refer to any nucleated blood cell that is not a nucleated erythrocyte. More specifically, leukocytes can be subdivided into two broad classes. The first class includes granulocytes, including, most prevalently, neutrophils, as well as eosinophils and basophils at low frequency. The second class, the non-granular or mononuclear leukocytes, includes monocytes and lymphocytes (e.g., T cells and B cells). There is an extensive literature in the art implicating leukocytes, e.g., neutrophils, monocytes and lymphocytes in a wide variety of disease processes, including inflammatory and rheumatic diseases, neurodegenerative diseases (such as Alzheimer's dementia), cardiovascular disease, endocrine diseases, transplant rejection, malignancy and infectious diseases, and other diseases listed in Table 1. Mononuclear cells are involved in the chronic immune response, while granulocytes, which make up approximately 60% of the leukocytes, have a non-specific and stereotyped response to acute inflammatory stimuli and often have a life span of only 24 hours.

In addition to their widespread involvement and/or implication in numerous disease related processes, leukocytes are particularly attractive substrates for clinical and experimental evaluation for a variety of reasons. Most importantly, they are

readily accessible at low cost from essentially every potential subject. Collection is minimally invasive and associated with little pain, disability or recovery time. Collection can be performed by minimally trained personnel (e.g., phlebotomists, medical technicians, etc.) in a variety of clinical and non-clinical settings without significant technological expenditure. Additionally, leukocytes are renewable, and thus available at multiple time points for a single subject.

Assembly of candidate libraries

At least two conceptually distinct approaches to the assembly of candidate libraries exist. Either, or both, or other, approaches can be favorably employed. The method of assembling, or identifying, candidate libraries is secondary to the criteria utilized for selecting appropriate library members. Most importantly, library members are assembled based on differential expression of RNA or protein products in leukocyte populations. More specifically, candidate nucleotide sequences are induced or suppressed, or expressed at increased or decreased levels in leukocytes from a subject with one or more disease or disease state (a disease criterion) relative to leukocytes from a subject lacking the specified disease criterion. Alternatively, or in addition, library members can be assembled from among nucleotide sequences that are differentially expressed in activated or resting leukocytes relative to other cell types.

Firstly, publication and sequence databases can be "mined" using a variety of search strategies, including, e.g., a variety of genomics and proteomics approaches. For example, currently available scientific and medical publication databases such as Medline, Current Contents, OMIM (online Mendelian inheritance in man) various Biological and Chemical Abstracts, Journal indexes, and the like can be searched using term or key-word searches, or by author, title, or other relevant search parameters. Many such databases are publicly available, and one of skill is well versed in strategies and procedures for identifying publications and their contents, e.g., genes, other nucleotide sequences, descriptions, indications, expression pattern, etc. Numerous databases are available through the internet for free or by subscription, *see*, e.g., http://www.ncbi.nlm.nih.gov/PubMed/; http://www3.infotrieve.com/; http://www.isinet.com/; http://www.sciencemag.org/. Additional or alternative publication or citation databases are also available that provide identical or similar types of information, any of which are favorable employed in the context of the invention. These databases can be searched for publications describing differential

gene expression in leukocytes between patient with and without diseases or conditions listed in Table 1. We identified the nucleotide sequences listed in Table 2 and some of the sequences listed in Table 8 (Example 20), using data mining methods.

Alternatively, a variety of publicly available and proprietary sequence databases (including GenBank, dbEST, UniGene, and TIGR and SAGE databases) including sequences corresponding to expressed nucleotide sequences, such as expressed sequence tags (ESTs) are available. For example, Genbank™ (http://www.ncbi.nlm.nih.gov/Genbank/) among others can be readily accessed and searched via the internet. These and other sequence and clone database resources are currently available; however, any number of additional or alternative databases comprising nucleotide sequence sequences, EST sequences, clone repositories, PCR primer sequences, and the like corresponding to individual nucleotide sequence sequences are also suitable for the purposes of the invention. Sequences from nucleotide sequences can be identified that are only found in libraries derived from leukocytes or sub-populations of leukocytes, for example see Table 2.

Alternatively, the representation, or relative frequency, of a nucleotide sequence may be determined in a leukocyte-derived nucleic acid library and compared to the representation of the sequence in non-leukocyte derived libraries. The representation of a nucleotide sequence correlates with the relative expression level of the nucleotide sequence in leukocytes and non-leukocytes. An oligonucleotide sequence which has increased or decreased representation in a leukocyte-derived nucleic acid library relative to a non-leukocyte-derived libraries is a candidate for a leukocyte-specific gene.

Nucleotide sequences identified as having specificity to activated or resting leukocytes or to leukocytes from patients or patient samples with a variety of disease types can be isolated for use in a candidate library for leukocyte expression profiling through a variety of mechanisms. These include, but are not limited to, the amplification of the nucleotide sequence from RNA or DNA using nucleotide sequence specific primers for PCR or RT-PCR, isolation of the nucleotide sequence using conventional cloning methods, the purchase of an IMAGE consortium cDNA clone (EST) with complimentary sequence or from the same expressed nucleotide sequence, design of oligonucleotides, preparation of synthetic nucleic acid sequence, or any other nucleic-acid based method. In addition, the protein product of the

nucleotide sequence can be isolated or prepared, and represented in a candidate library, using standard methods in the art, as described further below.

While the above discussion related primarily to "genomics" approaches, it is appreciated that numerous, analogous "proteomics" approaches are suitable to the present invention. For example, a differentially expressed protein product can, for example, be detected using western analysis, two-dimensional gel analysis, chromatographic separation, mass spectrometric detection, protein-fusion reporter constructs, colorometric assays, binding to a protein array, or by characterization of polysomal mRNA. The protein is further characterized and the nucleotide sequence encoding the protein is identified using standard techniques, e.g. by screening a cDNA library using a probe based on protein sequence information.

The second approach involves the construction of a differential expression library by any of a variety of means. Any one or more of differential screening, differential display or subtractive hybridization procedures, or other techniques that preferentially identify, isolate or amplify differentially expressed nucleotide sequences can be employed to produce a library of differentially expressed candidate nucleotide sequences, a subset of such a library, a partial library, or the like. Such methods are well known in the art. For example, peripheral blood leukocytes, (i.e., a mixed population including lymphocytes, monocytes and neutrophils), from multiple donor samples are pooled to prevent bias due to a single-donor's unique genotype. The pooled leukocytes are cultured in standard medium and stimulated with individual cytokines or growth factors e.g., with IL-2, IL-1, MCP1, TNFα, and/or IL8 according to well known procedures (*see*, e.g., Tough et al. (1999) ; Winston et al. (1999); Hansson et al. (1989)). Typically, leukocytes are recovered from Buffy coat preparations produced by centrifugation of whole blood. Alternatively, mononuclear cells (monocytes and lymphocytes) can be obtained by density gradient centrifugation of whole blood, or specific cell types (such as a T lymphocyte) can be isolated using affinity reagents to cell specific surface markers. Leukocytes may also be stimulated by incubation with ionomycin, and phorbol myristate acetate (PMA). This stimulation protocol is intended to non-specifically mimic "activation" of numerous pathways due to variety of disease conditions rather than to simulate any single disease condition or paradigm.

Using well known subtractive hybridization procedures (as described in, e.g., US Patent Numbers 5,958,738; 5589,339; 5,827,658; 5,712,127; 5,643,761) a library

is produced that is enriched for RNA species (messages) that are differentially expressed between test and control leukocyte populations. In some embodiments, the test population of leukocytes are simply stimulated as described above to emulate non-specific activation events, while in other embodiments the test population can be selected from subjects (or patients) with a specified disease or class of diseases. Typically, the control leukocyte population lacks the defining test condition, e.g., stimulation, disease state, diagnosis, genotype, etc. Alternatively, the total RNA from control and test leukocyte populations are prepared by established techniques, treated with DNAseI, and selected for messenger RNA with an intact 3′ end (i.e., polyA(+) messenger RNA) e.g., using commercially available kits according to the manufacturer's instructions e.g. Clontech. Double stranded cDNA is synthesized utilizing reverse transcriptase. Double stranded cDNA is then cut with a first restriction enzyme (e.g., *NlaIII*, that cuts at the recognition site: CATG, and cuts the cDNA sequence at approximately 256 bp intervals) that cuts the cDNA molecules into conveniently sized fragments.

The cDNAs prepared from the test population of leukocytes are divided into (typically 2) "tester" pools, while cDNAs prepared from the control population of leukocytes are designated the "driver" pool. Typically, pooled populations of cells from multiple individual donors are utilized and in the case of stimulated versus unstimulated cells, the corresponding tester and driver pools for any single subtraction reaction are derived from the same donor pool.

A unique double-stranded adapter is ligated to each of the tester cDNA populations using unphosphorylated primers so that only the sense strand is covalently linked to the adapter. An initial hybridization is performed consisting of each of the tester pools of cDNA (each with its corresponding adapter) and an excess of the driver cDNA. Typically, an excess of about 10-100 fold driver relative to tester is employed, although significantly lower or higher ratios can be empirically determined to provide more favorable results. The initial hybridization results in an initial normalization of the cDNAs such that high and low abundance messages become more equally represented following hybridization due to a failure of driver/tester hybrids to amplify.

A second hybridization involves pooling un-hybridized sequences from initial hybridizations together with the addition of supplemental driver cDNA. In this step, the expressed sequences enriched in the two tester pools following the initial

hybridization can hybridize. Hybrids resulting from the hybridization between members of each of the two tester pools are then recovered by amplification in a polymerase chain reaction (PCR) using primers specific for the unique adapters. Again, sequences originating in a tester pool that form hybrids with components of the driver pool are not amplified. Hybrids resulting between members of the same tester pool are eliminated by the formation of "panhandles" between their common 5' and 3' ends. For additional details, see, e.g., Lukyanov et al. (1997) Biochem Biophys Res Commun 230:285-8.

Typically, the tester and driver pools are designated in the alternative, such that the hybridization is performed in both directions to ensure recovery of messenger RNAs that are differentially expressed in either a positive or negative manner (i.e., that are turned on or turned off, up-regulated or down-regulated). Accordingly, it will be understood that the designation of test and control populations is to some extent arbitrary, and that a test population can just as easily be compared to leukocytes derived from a patient with the same of another disease of interest.

If so desired, the efficacy of the process can be assessed by such techniques as semi-quantitative PCR of known (i.e., control) nucleotide sequences, of varying abundance such as ß-actin. The resulting PCR products representing partial cDNAs of differentially expressed nucleotide sequences are then cloned (i.e., ligated) into an appropriate vector (e.g., a commercially available TA cloning vector, such as pGEM from Promega) and, optionally, transformed into competent bacteria for selection and screening.

Either of the above approaches, or both in combination, or indeed, any procedure, which permits the assembly of a collection of nucleotide sequences that are expressed in leukocytes, is favorably employed to produce the libraries of candidates useful for the identification of diagnostic nucleotide sets and disease specific target nucleotides of the invention. Additionally, any method that permits the assembly of a collection of nucleotides that are expressed in leukocytes and preferentially associated with one or more disease or condition, whether or not the nucleotide sequences are differentially expressed, is favorably employed in the context of the invention. Typically, libraries of about 2,000-10,000 members are produced (although libraries in excess of 10,000 are not uncommon). Following additional evaluation procedures, as described below, the proportion of unique clones in the candidate library can approximate 100%.

A candidate oligonucleotide sequence may be represented in a candidate library by a full-length or partial nucleic acid sequence, deoxyribonucleic acid (DNA) sequence, cDNA sequence, RNA sequence, synthetic oligonucleotides, etc. The nucleic acid sequence can be at least 19 nucleotides in length, at least 25 nucleotides, at least 40 nucleotides, at least 100 nucleotides, or larger. Alternatively, the protein product of a candidate nucleotide sequence may be represented in a candidate library using standard methods, as further described below.

Characterization of candidate oligonucleotide sequences

The sequence of individual members (e.g., clones, partial sequence listing in a database such as an EST, etc.) of the candidate oligonucleotide libraries is then determined by conventional sequencing methods well known in the art, e.g., by the dideoxy-chain termination method of Sanger et al. (1977) Proc Natl Acad Sci USA 74:5463-7; by chemical procedures, e.g., Maxam and Gilbert (1977) Proc Natl Acad Sci USA 74:560-4; or by polymerase chain reaction cycle sequencing methods, e.g., Olsen and Eckstein (1989) Nuc Acid Res 17:9613-20, DNA chip based sequencing techniques or variations, including automated variations (e.g., as described in Hunkapiller et al. (1991) Science 254:59-67; Pease et al. (1994) Proc Natl Acad Sci USA 91:5022-6), thereof. Numerous kits for performing the above procedures are commercially available and well known to those of skill in the art. Character strings corresponding to the resulting nucleotide sequences are then recorded (i.e., stored) in a database. Most commonly the character strings are recorded on a computer readable medium for processing by a computational device.

Generally, to facilitate subsequent analysis, a custom algorithm is employed to query existing databases in an ongoing fashion, to determine the identity, expression pattern and potential function of the particular members of a candidate library. The sequence is first processed, by removing low quality sequence. Next the vector sequences are identified and removed and sequence repeats are identified and masked. The remaining sequence is then used in a Blast algorithm against multiple publicly available, and/or proprietary databases, e.g., NCBI nucleotide, EST and protein databases, Unigene, and Human Genome Sequence. Sequences are also compared to all previously sequenced members of the candidate libraries to detect redundancy.

In some cases, sequences are of high quality, but do not match any sequence in the NCBI nr, human EST or Unigene databases. In this case the sequence is queried against the human genomic sequence. If a single chromosomal site is matched with a

high degree of confidence, that region of genomic DNA is identified and subjected to further analysis with a gene prediction program such as GRAIL. This analysis may lead to the identification of a new gene in the genomic sequence. This sequence can then be translated to identify the protein sequence that is encoded and that sequence can be further analyzed using tools such as Pfam, Blast P, or other protein structure prediction programs, as illustrated in Table 7. Typically, the above analysis is directed towards the identification of putative coding regions, e.g., previously unidentified open reading frames, confirming the presence of known coding sequences, and determining structural motifs or sequence similarities of the predicted protein (i.e., the conceptual translation product) in relation to known sequences. In addition, it has become increasingly possible to assemble "virtual cDNAs" containing large portions of coding region, simply through the assembly of available expressed sequence tags (ESTs). In turn, these extended nucleic acid and amino acid sequences allow the rapid expansion of substrate sequences for homology searches and structural and functional motif characterization. The results of these analysis permits the categorization of sequences according to structural characteristics, e.g., as structural proteins, proteins involved in signal transduction, cell surface or secreted proteins etc.

It is understood that full-length nucleotide sequences may also be identified using conventional methods, for example, library screening, RT-PCR, chromosome walking, etc., as described in *Sambrook and Ausebel, infra.*

Candidate nucleotide library of the invention

We identified members of a candidate nucleotide library that are differentially expressed in activated leukocytes and resting leukocytes. Accordingly, the invention provides the candidate leukocyte nucleotide library comprising the nucleotide sequences listed in Table 2, Table 3 and in the sequence listing. In another embodiment, the invention provides a candidate library comprising at least two nucleotide sequences listed in Table 2, Table 3, and the sequence listing. In another embodiment, the at least two nucleotide sequence are at least 19 nucleotides in length, at least 35 nucleotides, at least 40 nucleotides or at least 100 nucleotides. In some embodiments, the nucleotide sequences comprises deoxyribonucleic acid (DNA) sequence, ribonucleic acid (RNA) sequence, synthetic oligonucleotide sequence, or genomic DNA sequence. It is understood that the nucleotide sequences may each

correspond to one gene, or that several nucleotide sequences may correspond to one gene, or both.

The invention also provides probes to the candidate nucleotide library. In one embodiment of the invention, the probes comprise at least two nucleotide sequences listed in Table 2, Table 3, or the sequence listing which are differentially expressed in leukocytes in an individual with a least one disease criterion for at least one leukocyte-related disease and in leukocytes in an individual without the at least one disease criterion, wherein expression of the two or more nucleotide sequences is correlated with at least one disease criterion. It is understood that a probe may detect either the RNA expression or protein product expression of the candidate nucleotide library. Alternatively, or in addition, a probe can detect a genotype associated with a candidate nucleotide sequence, as further described below. In another embodiment, the probes for the candidate nucleotide library are immobilized on an array.

The candidate nucleotide library of the invention is useful in identifying diagnostic nucleotide sets of the invention, as described below. The candidate nucleotide sequences may be further characterized, and may be identified as a disease target nucleotide sequence and/or a novel nucleotide sequence, as described below. The candidate nucleotide sequences may also be suitable for use as imaging reagents, as described below.

Generation of Expression Patterns

RNA, DNA or protein sample procurement

Following identification or assembly of a library of differentially expressed candidate nucleotide sequences, leukocyte expression profiles corresponding to multiple members of the candidate library are obtained. Leukocyte samples from one or more subjects are obtained by standard methods. Most typically, these methods involve trans-cutaneous venous sampling of peripheral blood. While sampling of circulating leukocytes from whole blood from the peripheral vasculature is generally the simplest, least invasive, and lowest cost alternative, it will be appreciated that numerous alternative sampling procedures exist, and are favorably employed in some circumstances. No pertinent distinction exists, in fact, between leukocytes sampled from the peripheral vasculature, and those obtained, e.g., from a central line, from a central artery, or indeed from a cardiac catheter, or during a surgical procedure which accesses the central vasculature. In addition, other body fluids and tissues that are, at

least in part, composed of leukocytes are also desirable leukocyte samples. For example, fluid samples obtained from the lung during bronchoscopy may be rich in leukocytes, and amenable to expression profiling in the context of the invention, e.g., for the diagnosis, prognosis, or monitoring of lung transplant rejection, inflammatory lung diseases or infectious lung disease. Fluid samples from other tissues, e.g., obtained by endoscopy of the colon, sinuses, esophagus, stomach, small bowel, pancreatic duct, biliary tree, bladder, ureter, vagina, cervix or uterus, etc., are also suitable. Samples may also be obtained other sources containing leukocytes, e.g., from urine, bile, cerebrospinal fluid, feces, gastric or intestinal secretions, semen, or solid organ or joint biopsies.

Most frequently, mixed populations of leukocytes, such as are found in whole blood are utilized in the methods of the present invention. A crude separation, e.g., of mixed leukocytes from red blood cells, and/or concentration, e.g., over a sucrose, percoll or ficoll gradient, or by other methods known in the art, can be employed to facilitate the recovery of RNA or protein expression products at sufficient concentrations, and to reduce non-specific background. In some instances, it can be desirable to purify sub-populations of leukocytes, and methods for doing so, such as density or affinity gradients, flow cytometry, fluorescence Activated Cell Sorting (FACS), immuno-magnetic separation, "panning," and the like, are described in the available literature and below.

Obtaining DNA, RNA and protein samples for expression profiling

Expression patterns can be evaluated at the level of DNA, or RNA or protein products. For example, a variety of techniques are available for the isolation of RNA from whole blood. Any technique that allows isolation of mRNA from cells (in the presence or absence of rRNA and tRNA) can be utilized. In brief, one method that allows reliable isolation of total RNA suitable for subsequent gene expression analysis, is described as follows. Peripheral blood (either venous or arterial) is drawn from a subject, into one or more sterile, endotoxin free, tubes containing an anticoagulant (e.g., EDTA, citrate, heparin, etc.). Typically, the sample is divided into at least two portions. One portion, e.g., of 5-8 ml of whole blood is frozen and stored for future analysis, e.g., of DNA or protein. A second portion, e.g., of approximately 8 ml whole blood is processed for isolation of total RNA by any of a

variety of techniques as described in, e.g, Sambook, Ausubel, below, as well as U.S. Patent Numbers: 5,728,822 and 4,843,155.

Typically, a subject sample of mononuclear leukocytes obtained from about 8 ml of whole blood, a quantity readily available from an adult human subject under most circumstances, yields 5-20 μg of total RNA. This amount is ample, e.g., for labeling and hybridization to at least two probe arrays. Labeled probes for analysis of expression patterns of nucleotides of the candidate libraries are prepared from the subject's sample of RNA using standard methods. In many cases, cDNA is synthesized from total RNA using a polyT primer and labeled, e.g., radioactive or fluorescent, nucleotides. The resulting labeled cDNA is then hybridized to probes corresponding to members of the candidate nucleotide library, and expression data is obtained for each nucleotide sequence in the library. RNA isolated from subject samples (e.g., peripheral blood leukocytes, or leukocytes obtained from other biological fluids and samples) is next used for analysis of expression patterns of nucleotides of the candidate libraries.

In some cases, however, the amount of RNA that is extracted from the leukocyte sample is limiting, and amplification of the RNA is desirable. Amplification may be accomplished by increasing the efficiency of probe labeling, or by amplifying the RNA sample prior to labeling. It is appreciated that care must be taken to select an amplification procedure that does not introduce any bias (with respect to gene expression levels) during the amplification process.

Several methods are available that increase the signal from limiting amounts of RNA, e.g. use of the Clontech (Glass Fluorescent Labeling Kit) or Stratagene (Fairplay Microarray Labeling Kit), or the Micromax kit (New England Nuclear, Inc.). Alternatively, cDNA is synthesized from RNA using a T7- polyT primer, in the absence of label, and DNA dendrimers from Genisphere (3DNA Submicro) are hybridized to the poly T sequence on the primer, or to a different "capture sequence" which is complementary to a fluorescently labeled sequence. Each 3DNA molecule has 250 fluorescent molecules and therefore can strongly label each cDNA.

Alternatively, the RNA sample is amplified prior to labeling. For example, linear amplification may be performed, as described in U.S. Patent No. 6,132,997. A T7-polyT primer is used to generate the cDNA copy of the RNA. A second DNA strand is then made to complete the substrate for amplification. The T7 promoter

incorporated into the primer is used by a T7 polymerase to produce numerous antisense copies of the original RNA. Fluorescent dye labeled nucleotides are directly incorporated into the RNA. Alternatively, amino allyl labeled nucleotides are incorporated into the RNA, and then fluorescent dyes are chemically coupled to the amino allyl groups, as described in Hughes. Other exemplary methods for amplification are described below.

It is appreciated that the RNA isolated must contain RNA derived from leukocytes, but may also contain RNA from other cell types to a variable degree. Additionally, the isolated RNA may come from subsets of leukocytes, e.g. monocytes and/or T-lymphocytes, as described above. Such consideration of cell type used for the derivation of RNA depend on the method of expression profiling used.

DNA samples may be obtained for analysis of the presence of DNA mutations, single nucleotide polymorphisms (SNPs), or other polymorphisms. DNA is isolated using standard techniques, e.g. *Maniatus, supra.*

Expression of products of candidate nucleotides may also be assessed using proteomics. Protein(s) are detected in samples of patient serum or from leukocyte cellular protein. Serum is prepared by centrifugation of whole blood, using standard methods. Proteins present in the serum may have been produced from any of a variety of leukocytes and non-leukocyte cells, and include secreted proteins from leukocytes. Alternatively, leukocytes or a desired sub-population of leukocytes are prepared as described above. Cellular protein is prepared from leukocyte samples using methods well known in the art, e.g., Trizol (Invitrogen Life Technologies, cat # 15596108; Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156; Simms, D., Cizdziel, P.E., and Chomczynski, P. (1993) Focus® 15, 99; Chomczynski, P., Bowers-Finn, R., and Sabatini, L. (1987) J. of NIH Res. 6, 83; Chomczynski, P. (1993) Bio/Techniques 15, 532; Bracete, A.M., Fox, D.K., and Simms, D. (1998) Focus 20, 82; Sewall, A. and McRae, S. (1998) Focus 20, 36; Anal Biochem 1984 Apr;138(1):141-3, A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids; Wessel D, Flugge UI. (1984) Anal Biochem. 1984 Apr;138(1):141-143.

Obtaining expression patterns

Expression patterns, or profiles, of a plurality of nucleotides corresponding to members of the candidate library are then evaluated in one or more samples of leukocytes. Typically, the leukocytes are derived from patient peripheral blood

samples, although, as indicated above, many other sample sources are also suitable. These expression patterns constitute a set of relative or absolute expression values for a some number of RNAs or protein products corresponding to the plurality of nucleotide sequences evaluated, which is referred to herein as the subject's "expression profile" for those nucleotide sequences. While expression patterns for as few as one independent member of the candidate library can be obtained, it is generally preferable to obtain expression patterns corresponding to a larger number of nucleotide sequences, e.g., about 2, about 5, about 10, about 20, about 50, about 100, about 200, about 500, or about 1000, or more. The expression pattern for each differentially expressed component member of the library provides a finite specificity and sensitivity with respect to predictive value, e.g., for diagnosis, prognosis, monitoring, and the like.

Clinical Studies, Data and Patient Groups

For the purpose of discussion, the term subject, or subject sample of leukocytes, refers to an individual regardless of health and/or disease status. A subject can be a patient, a study participant, a control subject, a screening subject, or any other class of individual from whom a leukocyte sample is obtained and assessed in the context of the invention. Accordingly, a subject can be diagnosed with a disease, can present with one or more symptom of a disease, or a predisposing factor, such as a family (genetic) or medical history (medical) factor, for a disease, or the like. Alternatively, a subject can be healthy with respect to any of the aforementioned factors or criteria. It will be appreciated that the term "healthy" as used herein, is relative to a specified disease, or disease factor, or disease criterion, as the term "healthy" cannot be defined to correspond to any absolute evaluation or status. Thus, an individual defined as healthy with reference to any specified disease or disease criterion, can in fact be diagnosed with any other one or more disease, or exhibit any other one or more disease criterion.

Furthermore, while the discussion of the invention focuses, and is exemplified using human sequences and samples, the invention is equally applicable, through construction or selection of appropriate candidate libraries, to non-human animals, such as laboratory animals, e.g., mice, rats, guinea pigs, rabbits; domesticated livestock, e.g., cows, horses, goats, sheep, chicken, etc.; and companion animals, e.g., dogs, cats, etc.

Methods for obtaining expression data

Numerous methods for obtaining expression data are known, and any one or more of these techniques, singly or in combination, are suitable for determining expression profiles in the context of the present invention. For example, expression patterns can be evaluated by northern analysis, PCR, RT-PCR, Taq Man analysis, FRET detection, monitoring one or more molecular beacon, hybridization to an oligonucleotide array, hybridization to a cDNA array, hybridization to a polynucleotide array, hybridization to a liquid microarray, hybridization to a microelectric array, molecular beacons, cDNA sequencing, clone hybridization, cDNA fragment fingerprinting, serial analysis of gene expression (SAGE), subtractive hybridization, differential display and/or differential screening (*see*, e.g., Lockhart and Winzeler (2000) Nature 405:827-836, and references cited therein).

For example, specific PCR primers are designed to a member(s) of a candidate nucleotide library. cDNA is prepared from subject sample RNA by reverse transcription from a poly-dT oligonucleotide primer, and subjected to PCR. Double stranded cDNA may be prepared using primers suitable for reverse transcription of the PCR product, followed by amplification of the cDNA using in vitro transcription. The product of in vitro transcription is a sense-RNA corresponding to the original member(s) of the candidate library. PCR product may be also be evaluated in a number of ways known in the art, including real-time assessment using detection of labeled primers, e.g. TaqMan or molecular beacon probes. Technology platforms suitable for analysis of PCR products include the ABI 7700, 5700, or 7000 Sequence Detection Systems (Applied Biosystems, Foster City, CA), the MJ Research Opticon (MJ Research, Waltham, MA), the Roche Light Cycler (Roche Diagnositics, Indianapolis, IN), the Stratagene MX4000 (Stratagene, La Jolla, CA), and the Bio-Rad iCycler (Bio-Rad Laboratories, Hercules, CA). Alternatively, molecular beacons are used to detect presence of a nucleic acid sequence in an unamplified RNA or cDNA sample, or following amplification of the sequence using any method, e.g. IVT (In Vitro transcription) or NASBA (nucleic acid sequence based amplification). Molecular beacons are designed with sequences complementary to member(s) of a candidate nucleotide library, and are linked to fluorescent labels. Each probe has a different fluorescent label with non-overlapping emission wavelengths. For example,

expression of ten genes may be assessed using ten different sequence-specific molecular beacons.

Alternatively, or in addition, molecular beacons are used to assess expression of multiple nucleotide sequences at once. Molecular beacons with sequence complimentary to the members of a diagnostic nucleotide set are designed and linked to fluorescent labels. Each fluorescent label used must have a non-overlapping emission wavelength. For example, 10 nucleotide sequences can be assessed by hybridizing 10 sequence specific molecular beacons (each labeled with a different fluorescent molecule) to an amplified or un-amplified RNA or cDNA sample. Such an assay bypasses the need for sample labeling procedures.

Alternatively, or in addition bead arrays can be used to assess expression of multiple sequences at once. See, e.g, LabMAP 100, Luminex Corp, Austin, Texas). Alternatively, or in addition electric arrays are used to assess expression of multiple sequences, as exemplified by the e-Sensor technology of Motorola (Chicago, Ill.) or Nanochip technology of Nanogen (San Diego, CA.)

Of course, the particular method elected will be dependent on such factors as quantity of RNA recovered, practitioner preference, available reagents and equipment, detectors, and the like. Typically, however, the elected method(s) will be appropriate for processing the number of samples and probes of interest. Methods for high-throughput expression analysis are discussed below.

Alternatively, expression at the level of protein products of gene expression is performed. For example, protein expression, in a sample of leukocytes, can be evaluated by one or more method selected from among: western analysis, two-dimensional gel analysis, chromatographic separation, mass spectrometric detection, protein-fusion reporter constructs, colorimetric assays, binding to a protein array and characterization of polysomal mRNA. One particularly favorable approach involves binding of labeled protein expression products to an array of antibodies specific for members of the candidate library. Methods for producing and evaluating antibodies are widespread in the art, *see*, e.g., Coligan, *supra*; and Harlow and Lane (1989) Antibodies: A Laboratory Manual, Cold Spring Harbor Press, NY ("Harlow and Lane"). Additional details regarding a variety of immunological and immunoassay procedures adaptable to the present invention by selection of antibody reagents specific for the products of candidate nucleotide sequences can be found in, e.g., Stites and Terr (eds.)(1991) Basic and Clinical Immunology, 7th ed., and Paul, *supra.*

Another approach uses systems for performing desorption spectrometry. Commercially available systems, e.g., from Ciphergen Biosystems, Inc. (Fremont, CA) are particularly well suited to quantitative analysis of protein expression. Indeed, Protein Chip® arrays (*see*, e.g., http://www.ciphergen.com/) used in desorption spectrometry approaches provide arrays for detection of protein expression. Alternatively, affinity reagents, e.g., antibodies, small molecules, etc.) are developed that recognize epitopes of the protein product. Affinity assays are used in protein array assays, e.g. to detect the presence or absence of particular proteins. Alternatively, affinity reagents are used to detect expression using the methods described above. In the case of a protein that is expressed on the cell surface of leukocytes, labeled affinity reagents are bound to populations of leukocytes, and leukocytes expressing the protein are identified and counted using fluorescent activated cell sorting (FACS).

It is appreciated that the methods of expression evaluation discussed herein, although discussed in the context of discovery of diagnostic nucleotide sets, are equally applicable for expression evaluation when using diagnostic nucleotide sets for, e.g. diagnosis of diseases, as further discussed below.

High Throughput Expression Assays

A number of suitable high throughput formats exist for evaluating gene expression. Typically, the term high throughput refers to a format that performs at least about 100 assays, or at least about 500 assays, or at least about 1000 assays, or at least about 5000 assays, or at least about 10,000 assays, or more per day. When enumerating assays, either the number of samples or the number of candidate nucleotide sequences evaluated can be considered. For example, a northern analysis of, e.g., about 100 samples performed in a gridded array, e.g., a dot blot, using a single probe corresponding to a candidate nucleotide sequence can be considered a high throughput assay. More typically, however, such an assay is performed as a series of duplicate blots, each evaluated with a distinct probe corresponding to a different member of the candidate library. Alternatively, methods that simultaneously evaluate expression of about 100 or more candidate nucleotide sequences in one or more samples, or in multiple samples, are considered high throughput.

Numerous technological platforms for performing high throughput expression analysis are known. Generally, such methods involve a logical or physical array of

either the subject samples, or the candidate library, or both. Common array formats include both liquid and solid phase arrays. For example, assays employing liquid phase arrays, e.g., for hybridization of nucleic acids, binding of antibodies or other receptors to ligand, etc., can be performed in multiwell, or microtiter, plates. Microtiter plates with 96, 384 or 1536 wells are widely available, and even higher numbers of wells, e.g, 3456 and 9600 can be used. In general, the choice of microtiter plates is determined by the methods and equipment, e.g., robotic handling and loading systems, used for sample preparation and analysis. Exemplary systems include, e.g., the ORCA™ system from Beckman-Coulter, Inc. (Fullerton, CA) and the Zymate systems from Zymark Corporation (Hopkinton, MA).

Alternatively, a variety of solid phase arrays can favorably be employed in to determine expression patterns in the context of the invention. Exemplary formats include membrane or filter arrays (e.g, nitrocellulose, nylon), pin arrays, and bead arrays (e.g., in a liquid "slurry"). Typically, probes corresponding to nucleic acid or protein reagents that specifically interact with (e.g., hybridize to or bind to) an expression product corresponding to a member of the candidate library are immobilized, for example by direct or indirect cross-linking, to the solid support. Essentially any solid support capable of withstanding the reagents and conditions necessary for performing the particular expression assay can be utilized. For example, functionalized glass, silicon, silicon dioxide, modified silicon, any of a variety of polymers, such as (poly)tetrafluoroethylene, (poly)vinylidenedifluoride, polystyrene, polycarbonate, or combinations thereof can all serve as the substrate for a solid phase array.

In a preferred embodiment, the array is a "chip" composed, e.g., of one of the above specified materials. Polynucleotide probes, e.g., RNA or DNA, such as cDNA, synthetic oligonucleotides, and the like, or binding proteins such as antibodies, that specifically interact with expression products of individual components of the candidate library are affixed to the chip in a logically ordered manner, i.e., in an array. In addition, any molecule with a specific affinity for either the sense or anti-sense sequence of the marker nucleotide sequence (depending on the design of the sample labeling), can be fixed to the array surface without loss of specific affinity for the marker and can be obtained and produced for array production, for example, proteins that specifically recognize the specific nucleic acid sequence of the marker,

ribozymes, peptide nucleic acids (PNA), or other chemicals or molecules with specific affinity.

Detailed discussion of methods for linking nucleic acids and proteins to a chip substrate, are found in, e.g., US Patent No. 5,143,854 "LARGE SCALE PHOTOLITHOGRAPHIC SOLID PHASE SYNTHESIS OF POLYPEPTIDES AND RECEPTOR BINDING SCREENING THEREOF" to Pirrung et al., issued, September 1, 1992; US Patent No. 5,837,832 "ARRAYS OF NUCLEIC ACID PROBES ON BIOLOGICAL CHIPS" to Chee et al., issued November 17, 1998; US Patent No. 6,087,112 "ARRAYS WITH MODIFIED OLIGONUCLEOTIDE AND POLYNUCLEOTIDE COMPOSITIONS" to Dale, issued July 11, 2000; US Patent No. 5,215,882 "METHOD OF IMMOBILIZING NUCLEIC ACID ON A SOLID SUBSTRATE FOR USE IN NUCLEIC ACID HYBRIDIZATION ASSAYS" to Bahl et al., issued June 1, 1993; US Patent No. 5,707,807 "MOLECULAR INDEXING FOR EXPRESSED GENE ANALYSIS" to Kato, issued January 13, 1998; US Patent No. 5,807,522 "METHODS FOR FABRICATING MICROARRAYS OF BIOLOGICAL SAMPLES" to Brown et al., issued September 15, 1998; US Patent No. 5,958,342 "JET DROPLET DEVICE" to Gamble et al., issued Sept. 28, 1999; US Patent 5,994,076 "METHODS OF ASSAYING DIFFERENTIAL EXPRESSION" to Chenchik et al., issued Nov. 30, 1999; US Patent No. 6,004,755 "QUANTITATIVE MICROARRAY HYBRIDIZATION ASSAYS" to Wang, issued Dec. 21, 1999; US Patent No. 6,048,695 "CHEMICALLY MODIFIED NUCLEIC ACIDS AND METHOD FOR COUPLING NUCLEIC ACIDS TO SOLID SUPPORT" to Bradley et al., issued April 11, 2000; US Patent No. 6,060,240 "METHODS FOR MEASURING RELATIVE AMOUNTS OF NUCLEIC ACIDS IN A COMPLEX MIXTURE AND RETRIEVAL OF SPECIFIC SEQUENCES THEREFROM" to Kamb et al., issued May 9, 2000; US Patent No. 6,090,556 "METHOD FOR QUANTITATIVELY DETERMINING THE EXPRESSION OF A GENE" to Kato, issued July 18, 2000; and US Patent 6,040,138 "EXPRESSION MONITORING BY HYBRIDIZATION TO HIGH DENSITY OLIGONUCLEOTIDE ARRAYS" to Lockhart et al., issued March 21, 2000.

For example, cDNA inserts corresponding to candidate nucleotide sequences, in a standard TA cloning vector are amplified by a polymerase chain reaction for approximately 30-40 cycles. The amplified PCR products are then arrayed onto a glass support by any of a variety of well known techniques, e.g., the VSLIPS™

technology described in US Patent No. 5,143,854. RNA, or cDNA corresponding to RNA, isolated from a subject sample of leukocytes is labeled, e.g., with a fluorescent tag, and a solution containing the RNA (or cDNA) is incubated under conditions favorable for hybridization, with the "probe" chip. Following incubation, and washing to eliminate non-specific hybridization, the labeled nucleic acid bound to the chip is detected qualitatively or quantitatively, and the resulting expression profile for the corresponding candidate nucleotide sequences is recorded. It is appreciated that the probe used for diagnostic purposes may be identical to the probe used during diagnostic nucleotide sequence discovery and validation. Alternatively, the probe sequence may be different than the sequence used in diagnostic nucleotide sequence discovery and validation. Multiple cDNAs from a nucleotide sequence that are non-overlapping or partially overlapping may also be used.

In another approach, oligonucleotides corresponding to members of a candidate nucleotide library are synthesized and spotted onto an array. Alternatively, oligonucleotides are synthesized onto the array using methods known in the art, e.g. Hughes, et al. *supra*. The oligonucleotide is designed to be complementary to any portion of the candidate nucleotide sequence. In addition, in the context of expression analysis for, e.g. diagnostic use of diagnostic nucleotide sets, an oligonucleotide can be designed to exhibit particular hybridization characteristics, or to exhibit a particular specificity and/or sensitivity, as further described below.

Hybridization signal may be amplified using methods known in the art, and as described herein, for example use of the Clontech kit (Glass Fluorescent Labeling Kit), Stratagene kit (Fairplay Microarray Labeling Kit), the Micromax kit (New England Nuclear, Inc.), the Genisphere kit (3DNA Submicro), linear amplification, e.g. as described in U.S. Patent No. 6,132,997 or described in Hughes, TR, et al., Nature Biotechnology, 19:343-347 (2001) and/or Westin et al. Nat Biotech. 18:199-204.

Alternatively, fluorescently labeled cDNA are hybridized directly to the microarray using methods known in the art. For example, labeled cDNA are generated by reverse transcription using Cy3- and Cy5-conjugated deoxynucleotides, and the reaction products purified using standard methods. It is appreciated that the methods for signal amplification of expression data useful for identifying diagnostic nucleotide sets are also useful for amplification of expression data for diagnostic purposes.

Microarray expression may be detected by scanning the microarray with a variety of laser or CCD-based scanners, and extracting features with numerous software packages, for example, Imagene (Biodiscovery), Feature Extraction (Agilent), Scanalyze (Eisen, M. 1999. SCANALYZE User Manual; Stanford Univ., Stanford, CA. Ver 2.32.), GenePix (Axon Instruments).

In another approach, hybridization to microelectric arrays is performed, e.g. as described in Umek et al (2001) J Mol Diagn. 3:74-84. An affinity probe, e.g. DNA, is deposited on a metal surface. The metal surface underlying each probe is connected to a metal wire and electrical signal detection system. Unlabelled RNA or cDNA is hybridized to the array, or alternatively, RNA or cDNA sample is amplified before hybridization, e.g. by PCR. Specific hybridization of sample RNA or cDNA results in generation of an electrical signal, which is transmitted to a detector. See Westin (2000) Nat Biotech. 18:199-204 (describing anchored multiplex amplification of a microelectronic chip array); Edman (1997) NAR 25:4907-14; Vignali (2000) J Immunol Methods 243:243-55.

In another approach, a microfluidics chip is used for RNA sample preparation and analysis. This approach increases efficiency because sample preparation and analysis are streamlined. Briefly, microfluidics may be used to sort specific leukocyte sub-populations prior to RNA preparation and analysis. Microfluidics chips are also useful for, e.g., RNA preparation, and reactions involving RNA (reverse transcription, RT-PCR). Briefly, a small volume of whole, anti-coagulated blood is loaded onto a microfluidics chip, for example chips available from Caliper (Mountain View, CA) or Nanogen (San Diego, CA.) A microfluidics chip may contain channels and reservoirs in which cells are moved and reactions are performed. Mechanical, electrical, magnetic, gravitational, centrifugal or other forces are used to move the cells and to expose them to reagents. For example, cells of whole blood are moved into a chamber containing hypotonic saline, which results in selective lysis of red blood cells after a 20-minute incubation. Next, the remaining cells (leukocytes) are moved into a wash chamber and finally, moved into a chamber containing a lysis buffer such as guanidine isothyocyanate. The leukocyte cell lysate is further processed for RNA isolation in the chip, or is then removed for further processing, for example, RNA extraction by standard methods. Alternatively, the microfluidics chip is a circular disk containing ficoll or another density reagent. The blood sample is injected into the center of the disc, the disc is rotated at a speed that generates a

centrifugal force appropriate for density gradient separation of mononuclear cells, and the separated mononuclear cells are then harvested for further analysis or processing.

It is understood that the methods of expression evaluation, above, although discussed in the context of discovery of diagnostic nucleotide sets, are also applicable for expression evaluation when using diagnostic nucleotide sets for, e.g. diagnosis of diseases, as further discussed below.

Evaluation of expression patterns

Expression patterns can be evaluated by qualitative and/or quantitative measures. Certain of the above described techniques for evaluating gene expression (as RNA or protein products) yield data that are predominantly qualitative in nature. That is, the methods detect differences in expression that classify expression into distinct modes without providing significant information regarding quantitative aspects of expression. For example, a technique can be described as a qualitative technique if it detects the presence or absence of expression of a candidate nucleotide sequence, i.e., an on/off pattern of expression. Alternatively, a qualitative technique measures the presence (and/or absence) of different alleles, or variants, of a gene product.

In contrast, some methods provide data that characterizes expression in a quantitative manner. That is, the methods relate expression on a numerical scale, e.g., a scale of 0-5, a scale of 1-10, a scale of + - +++, from grade 1 to grade 5, a grade from a to z, or the like. It will be understood that the numerical, and symbolic examples provided are arbitrary, and that any graduated scale (or any symbolic representation of a graduated scale) can be employed in the context of the present invention to describe quantitative differences in nucleotide sequence expression. Typically, such methods yield information corresponding to a relative increase or decrease in expression.

Any method that yields either quantitative or qualitative expression data is suitable for evaluating expression of candidate nucleotide sequence in a subject sample of leukocytes. In some cases, e.g., when multiple methods are employed to determine expression patterns for a plurality of candidate nucleotide sequences, the recovered data, e.g., the expression profile, for the nucleotide sequences is a combination of quantitative and qualitative data.

In some applications, expression of the plurality of candidate nucleotide sequences is evaluated sequentially. This is typically the case for methods that can be characterized as low- to moderate-throughput. In contrast, as the throughput of the elected assay increases, expression for the plurality of candidate nucleotide sequences in a sample or multiple samples of leukocytes, is assayed simultaneously. Again, the methods (and throughput) are largely determined by the individual practitioner, although, typically, it is preferable to employ methods that permit rapid, e.g. automated or partially automated, preparation and detection, on a scale that is time-efficient and cost-effective.

It is understood that the preceding discussion, while directed at the assessment of expression of the members of candidate libraries, is also applies to the assessment of the expression of members of diagnostic nucleotide sets, as further discussed below.

Genotyping

In addition to, or in conjunction with the correlation of expression profiles and clinical data, it is often desirable to correlate expression patterns with the subject's genotype at one or more genetic loci. The selected loci can be, for example, chromosomal loci corresponding to one or more member of the candidate library, polymorphic alleles for marker loci, or alternative disease related loci (not contributing to the candidate library) known to be, or putatively associated with, a disease (or disease criterion). Indeed, it will be appreciated, that where a (polymorphic) allele at a locus is linked to a disease (or to a predisposition to a disease), the presence of the allele can itself be a disease criterion.

Numerous well known methods exist for evaluating the genotype of an individual, including southern analysis, restriction fragment length polymorphism (RFLP) analysis, polymerase chain reaction (PCR), amplification length polymorphism (AFLP) analysis, single stranded conformation polymorphism (SSCP) analysis, single nucleotide polymorphism (SNP) analysis (e.g., via PCR, Taqman or molecular beacons), among many other useful methods. Many such procedures are readily adaptable to high throughput and/or automated (or semi-automated) sample preparation and analysis methods. Most, can be performed on nucleic acid samples recovered via simple procedures from the same sample of leukocytes as yielded the

material for expression profiling. Exemplary techniques are described in, e.g., Sambrook, and Ausubel, *supra*.

Identification of the diagnostic nucleotide sets of the invention

Identification of diagnostic nucleotide sets and disease specific target nucleotide sequence proceeds by correlating the leukocyte expression profiles with data regarding the subject's health status to produce a data set designated a "molecular signature." Examples of data regarding a patient's health status, also termed "disease criteria(ion)", is described below and in the Section titled "selected diseases," below. Methods useful for correlation analysis are further described elsewhere in the specification.

Generally, relevant data regarding the subject's health status includes retrospective or prospective health data, e.g., in the form of the subject's medical history, as provided by the subject, physician or third party, such as, medical diagnoses, laboratory test results, diagnostic test results, clinical events, or medication lists, as further described below. Such data may include information regarding a patient's response to treatment and/or a particular medication and data regarding the presence of previously characterized "risk factors." For example, cigarette smoking and obesity are previously identified risk factors for heart disease. Further examples of health status information, including diseases and disease criteria, is described in the section titled Selected diseases, below.

Typically, the data describes prior events and evaluations (i.e., retrospective data). However, it is envisioned that data collected subsequent to the sampling (i.e., prospective data) can also be correlated with the expression profile. The tissue sampled, e.g., peripheral blood, bronchial lavage, etc., can be obtained at one or more multiple time points and subject data is considered retrospective or prospective with respect to the time of sample procurement.

Data collected at multiple time points, called "longitudinal data", is often useful, and thus, the invention encompasses the analysis of patient data collected from the same patient at different time points. Analysis of paired samples, such as samples from a patient at different time, allows identification of differences that are specifically related to the disease state since the genetic variability specific to the patient is controlled for by the comparison. Additionally, other variables that exist between patients may be controlled for in this way, for example, the presence or

absence of inflammatory diseases (e.g., rheumatoid arthritis) the use of medications that may effect leukocyte gene expression, the presence or absence of co-morbid conditions, etc. Methods for analysis of paired samples are further described below. Moreover, the analysis of a pattern of expression profiles (generated by collecting multiple expression profiles) provides information relating to changes in expression level over time, and may permit the determination of a rate of change, a trajectory, or an expression curve. Two longitudinal samples may provide information on the change in expression of a gene over time, while three longitudinal samples may be necessary to determine the "trajectory" of expression of a gene. Such information may be relevant to the diagnosis of a disease. For example, the expression of a gene may vary from individual to individual, but a clinical event, for example , a heart attack, may cause the level of expression to double in each patient. In this example, clinically interesting information is gleaned from the change in expression level, as opposed to the absolute level of expression in each individual.

Generally, small sample sizes of 10-40 samples from 10-20 individuals are used to identify a diagnostic nucleotide set. Larger sample sizes are generally necessary to validate the diagnostic nucleotide set for use in large and varied patient populations, as further described below. For example, extension of gene expression correlations to varied ethnic groups, demographic groups, nations, peoples or races may require expression correlation experiments on the population of interest.

Expression Reference Standards

Expression profiles derived from a patient (i.e., subjects diagnosed with, or exhibiting symptoms of, or exhibiting a disease criterion, or under a doctor's care for a disease) sample are compared to a control or standard expression RNA to facilitate comparison of expression profiles (e.g. of a set of candidate nucleotide sequences) from a group of patients relative to each other (i.e., from one patient in the group to other patients in the group, or to patients in another group).

For example, in one approach to identifying diagnostic nucleotide sets, expression profiles derived from patient samples are compared to a expression reference "standard." Standard expression reference can be, for example, RNA derived from resting cultured leukocytes or commercially available reference RNA, such as Universal reference RNA from Stratagene. *See* Nature, V406, 8-17-00, p. 747-752. Use of an expression reference standard is particularly useful when the expression of large numbers of nucleotide sequences is assayed, e.g. in an array, and

in certain other applications, e.g. qualitative PCR, RT-PCR, etc., where it is desirable to compare a sample profile to a standard profile, and/or when large numbers of expression profiles, e.g. a patient population, are to be compared. Generally, an expression reference standard should be available in large quantities, should be a good substrate for amplification and labeling reactions, and should be capable of detecting a large percentage of candidate nucleic acids using suitable expression profiling technology.

Alternatively, or in addition, the expression profile derived from a patient sample is compared with the expression of an internal reference control gene, for example, β-actin or CD4. The relative expression of the profiled genes and the internal reference control gene (from the same individual) is obtained. An internal reference control may also be used with a reference RNA. For example, an expression profile for "gene 1" and the gene encoding CD4 can be determined in a patient sample and in a reference RNA. The expression of each gene can be expressed as the "relative" ratio of expression the gene in the patient sample compared with expression of the gene in the reference RNA. The expression ratio (sample/reference) for gene 1 may be divided by the expression ration for CD4 (sample/reference) and thus the relative expression of gene 1 to CD4 is obtained.

The invention also provides a buffy coat control RNA useful for expression profiling, and a method of using control RNA produced from a population of buffy coat cells, the white blood cell layer derived from the centrifugation of whole blood. Buffy coat contains all white blood cells, including granulocytes, mononuclear cells and platelets. The invention also provides a method of preparing control RNA from buffy coat cells for use in expression profile analysis of leukocytes. Buffy coat fractions are obtained, e.g. from a blood bank or directly from individuals, preferably from a large number of individuals such that bias from individual samples is avoided and so that the RNA sample represents an average expression of a healthy population. Buffy coat fractions from about 50 or about 100, or more individuals are preferred. 10 ml buffy coat from each individual is used. Buffy coat samples are treated with an erthythrocyte lysis buffer, so that erthythrocytes are selectively removed. The leukocytes of the buffy coat layer are collected by centrifugation. Alternatively, the buffy cell sample can be further enriched for a particular leukocyte sub-populations, e.g. mononuclear cells, T-lymphocytes, etc. To enrich for mononuclear cells, the

buffy cell pellet, above, is diluted in PBS (phosphate buffered saline) and loaded onto a non-polystyrene tube containing a polysucrose and sodium diatrizoate solution adjusted to a density of 1.077+/-0.001 g/ml. To enrich for T-lymphocytes, 45 ml of whole blood is treated with RosetteSep (Stem Cell Technologies), and incubated at room temperature for 20 minutes. The mixture is diluted with an equal volume of PBS plus 2% FBS and mixed by inversion. 30 ml of diluted mixture is layered on top of 15 ml DML medium (Stem Cell Technologies). The tube is centrifuged at 1200 x g, and the enriched cell layer at the plasma : medium interface is removed, washed with PBS + 2% FBS, and cells collected by centrifugation at 1200 x g. The cell pellet is treated with 5 ml of erythrocyte lysis buffer (EL buffer, Qiagen) for 10 minutes on ice, and enriched T-lymphoctes are collected by centrifugation.

In addition or alternatively, the buffy cells (whole buffy coat or sub-population, e.g. mononuclear fraction) can be cultured *in vitro* and subjected to stimulation with cytokines or activating chemicals such as phorbol esters or ionomycin. Such stimuli may increase expression of nucleotide sequences that are expressed in activated immune cells and might be of interest for leukocyte expression profiling experiments.

Following sub-population selection and/or further treatment, e.g. stimulation as described above, RNA is prepared using standard methods. For example, cells are pelleted and lysed with a phenol/guanidinium thiocyanate and RNA is prepared. RNA can also be isolated using a silica gel-based purification column or the column method can be used on RNA isolated by the phenol/guanidinium thiocyanate method. RNA from individual buffy coat samples can be pooled during this process, so that the resulting reference RNA represents the RNA of many individuals and individual bias is minimized or eliminated. In addition, a new batch of buffy coat reference RNA can be directly compared to the last batch to ensure similar expression pattern from one batch to another, using methods of collecting and comparing expression profiles described above/below. One or more expression reference controls are used in an experiment. For example, RNA derived from one or more of the following sources can be used as controls for an experiment: stimulated or unstimulated whole buffy coat, stimulated or unstimulated peripheral mononuclear cells, or stimulated or unstimulated T-lymphocytes.

Alternatively, the expression reference standard can be derived from any subject or class of subjects including healthy subjects or subjects diagnosed with the

same or a different disease or disease criterion. Expression profiles from subjects in two distinct classes are compared to determine which subset of nucleotide sequences in the candidate library best distinguish between the two subject classes, as further discussed below. It will be appreciated that in the present context, the term "distinct classes" is relevant to at least one distinguishable criterion relevant to a disease of interest, a "disease criterion." The classes can, of course, demonstrate significant overlap (or identity) with respect to other disease criteria, or with respect to disease diagnoses, prognoses, or the like. The mode of discovery involves, e.g., comparing the molecular signature of different subject classes to each other (such as patient to control, patients with a first diagnosis to patients with a second diagnosis, etc.) or by comparing the molecular signatures of a single individual taken at different time points. The invention can be applied to a broad range of diseases, disease criteria, conditions and other clinical and/or epidemiological questions, as further discussed above/below.

It is appreciated that while the present discussion pertains to the use of expression reference controls while identifying diagnostic nucleotide sets, expression reference controls are also useful during use of diagnostic nucleotide sets, e.g. use of a diagnostic nucleotide set for diagnosis of a disease, as further described below.

Analysis of expression profiles

In order to facilitate ready access, e.g., for comparison, review, recovery, and/or modification, the molecular signatures/expression profiles are typically recorded in a database. Most typically, the database is a relational database accessible by a computational device, although other formats, e.g., manually accessible indexed files of expression profiles as photographs, analogue or digital imaging readouts, spreadsheets, etc. can be used. Further details regarding preferred embodiments are provided below. Regardless of whether the expression patterns initially recorded are analog or digital in nature and/or whether they represent quantitative or qualitative differences in expression, the expression patterns, expression profiles (collective expression patterns), and molecular signatures (correlated expression patterns) are stored digitally and accessed via a database. Typically, the database is compiled and maintained at a central facility, with access being available locally and/or remotely.

As additional samples are obtained, and their expression profiles determined and correlated with relevant subject data, the ensuing molecular signatures are likewise recorded in the database. However, rather than each subsequent addition

being added in an essentially passive manner in which the data from one sample has little relation to data from a second (prior or subsequent) sample, the algorithms optionally additionally query additional samples against the existing database to further refine the association between a molecular signature and disease criterion. Furthermore, the data set comprising the one (or more) molecular signatures is optionally queried against an expanding set of additional or other disease criteria. The use of the database in integrated systems and web embodiments is further described below.

Analysis of expression profile data from arrays

Expression data is analyzed using methods well known in the art, including the software packages Imagene (Biodiscovery, Marina del Rey, CA), Feature Extraction (Agilent, Palo Alto, CA), and Scanalyze (Stanford University). In the discussion that follows, a "feature" refers to an individual spot of DNA on an array . Each gene may have more than one feature. For example, hybridized microarrays are scanned and analyzed on an Axon Instruments scanner using GenePix 3.0 software (Axon Instruments, Union City, CA). The data extracted by GenePix is used for all downstream quality control and expression evaluation. The data is derived as follows. The data for all features flagged as "not found" by the software is removed from the dataset for individual hybridizations. The "not found" flag by GenePix indicates that the software was unable to discriminate the feature from the background. Each feature is examined to determine the value of its signal. The median pixel intensity of the background (B_n) is subtracted from the median pixel intensity of the feature (F_n) to produce the background-subtracted signal (hereinafter, "BGSS"). The BGSS is divided by the standard deviation of the background pixels to provide the signal-to-noise ratio (hereinafter, "S/N"). Features with a S/N of three or greater in both the Cy3 channel (corresponding to the sample RNA) and Cy5 channel (corresponding to the reference RNA) are used for further analysis (hereinafter denoted "useable features"). Alternatively, different S/Ns are used for selecting expression data for an analysis. For example, only expression data with signal to noise ratios > 3 might be used in an analysis.

For each usable feature (*i*), the expression level (*e*) is expressed as the logarithm of the ratio (R) of the Background Subtracted Signal (hereinafter "BGSS") for the Cy3 (sample RNA) channel divided by the BGSS for the Cy5 channel (reference RNA). This "log ratio" value is used for comparison to other experiments.

$$R_i = \frac{BGSS_{sample}}{BGSS_{reference}} \quad (0.1)$$

$$e_i = \log r_i \quad (0.2)$$

Variation in signal across hybridizations may be caused by a number of factors affecting hybridization, DNA spotting, wash conditions, and labeling efficiency.

A single reference RNA may be used with all of the experimental RNAs, permitting multiple comparisons in addition to individual comparisons. By comparing sample RNAs to the same reference, the gene expression levels from each sample are compared across arrays, permitting the use of a consistent denominator for our experimental ratios.

Scaling

The data may be scaled (normalized) to control for labeling and hybridization variability within the experiment, using methods known in the art. Scaling is desirable because it facilitates the comparison of data between different experiments, patients, etc. Generally the BGSS are scaled to a factor such as the median, the mean, the trimmed mean, and percentile. Additional methods of scaling include: to scale between 0 and 1, to subtract the mean, or to subtract the median.

Scaling is also performed by comparison to expression patterns obtained using a common reference RNA, as described in greater detail above. As with other scaling methods, the reference RNA facilitates multiple comparisons of the expression data, e.g., between patients, between samples, etc. Use of a reference RNA provides a consistent denominator for experimental ratios.

In addition to the use of a reference RNA, individual expression levels may be adjusted to correct for differences in labeling efficiency between different hybridization experiments, allowing direct comparison between experiments with different overall signal intensities, for example. A scaling factor (a) may be used to adjust individual expression levels as follows. The median of the scaling factor (a), for example, BGSS, is determined for the set of all features with a S/N greater than three. Next, the $BGSS_i$ (the BGSS for each feature "i") is divided by the median for

all features (a), generating a scaled ratio. The scaled ration is used to determine the expression value for the feature (e_i), or the log ratio.

$$S_i = \frac{BGSS_i}{a} \tag{0.3}$$

$$e_i = \log\left(\frac{Cy3S_i}{Cy5S_i}\right) \tag{0.4}$$

In addition, or alternatively, control features are used to normalize the data for labeling and hybridization variability within the experiment. Control feature may be cDNA for genes from the plant, *Arabidopsis thaliana*, that are included when spotting the mini-array. Equal amounts of RNA complementary to control cDNAs are added to each of the samples before they were labeled. Using the signal from these control genes, a normalization constant (L) is determined according to the following formula:

$$L_j = \frac{\frac{\sum_{i=1}^{N} BGSS_{j,i}}{N}}{\frac{\sum_{j=1}^{K}\frac{\sum_{i=1}^{N} BGSS_{j,i}}{N}}{K}}$$

where $BGSS_i$ is the signal for a specific feature, N is the number of *A. thaliana* control features, K is the number of hybridizations, and L_j is the normalization constant for each individual hybridization.

Using the formula above, the mean for all control features of a particular hybridization and dye (e.g., Cy3) is calculated. The control feature means for all Cy3 hybridizations are averaged, and the control feature mean in one hybridization divided by the average of all hybridizations to generate a normalization constant for that particular Cy3 hybridization (L_j), which is used as a in equation (0.3). The same normalization steps may be performed for Cy3 and Cy5 values.

Many additional methods for normalization exist and can be applied to the data. In one method, the average ratio of Cy3 BGSS / Cy5 BGSS is determined for all features on an array. This ratio is then scaled to some arbitrary number, such as 1 or some other number. The ratio for each probe is then multiplied by the scaling

factor required to bring the average ratio to the chosen level. This is performed for each array in an analysis. Alternatively, the ratios are normalized to the average ratio across all arrays in an analysis.

Correlation analysis

Correlation analysis is performed to determine which array probes have expression behavior that best distinguishes or serves as markers for relevant groups of samples representing a particular clinical condition. Correlation analysis, or comparison among samples representing different disease criteria (e.g., clinical conditions), is performed using standard statistical methods. Numerous algorithms are useful for correlation analysis of expression data, and the selection of algorithms depends in part on the data analysis to be performed. For example, algorithms can be used to identify the single most informative gene with expression behavior that reliably classifies samples, or to identify all the genes useful to classify samples. Alternatively, algorithms can be applied that determine which set of 2 or more genes have collective expression behavior that accurately classifies samples. The use of multiple expression markers for diagnostics may overcome the variability in expression of a gene between individuals, or overcome the variability intrinsic to the assay. Multiple expression markers may include redundant markers, in that two or more genes or probes may provide the same information with respect to diagnosis. This may occur, for example, when two or more genes or gene probes are coordinately expressed. It will be appreciated that while the discussion above pertains to the analysis of RNA expression profiles the discussion is equally applicable to the analysis of profiles of proteins or other molecular markers.

Prior to analysis, expression profile data may be formatted or prepared for analysis using methods known in the art. For example, often the log ratio of scaled expression data for every array probe is calculated using the following formula:

log (Cy 3 BGSS/ Cy5 BGSS), where Cy 3 signal corresponds to the expression of the gene in the clinical sample, and Cy5 signal corresponds to expression of the gene in the reference RNA.

Data may be further filtered depending on the specific analysis to be done as noted below. For example, filtering may be aimed at selecting only samples with expression above a certain level, or probes with variability above a certain level between sample sets.

The following non-limiting discussion consider several statistical methods known in the art. Briefly, the t-test and ANOVA are used to identify single genes with expression differences between or among populations, respectively. Multivariate methods are used to identify a set of two or more genes for which expression discriminates between two disease states more specifically than expression of any single gene.

t-test

The simplest measure of a difference between two groups is the Student's t test. See, e.g., Welsh et al. (2001) Proc Natl Acad Sci USA 98:1176-81 (demonstrating the use of an unpaired Student's t-test for the discovery of differential gene expression in ovarian cancer samples and control tissue samples). The t- test assumes equal variance and normally distributed data. This test identifies the probability that there is a difference in expression of a single gene between two groups of samples. The number of samples within each group that is required to achieve statistical significance is dependent upon the variation among the samples within each group. The standard formula for a t-test is:

$$t(e_i) = \frac{\bar{e}_{i,c} - \bar{e}_{i,t}}{\sqrt{(s_{i,c}^2/n_c) + (s_{i,t}^2/n_t)}}, \qquad (0.5)$$

where $\bar{e}_i$ is the difference between the mean expression level of gene i in groups c and t, $s_{i,c}$ is the variance of gene x in group c and $s_{i,t}$ is the variance of gene x in group t. n_c and n_t are the numbers of samples in groups c and t.

The combination of the t statistic and the degrees of freedom [$\min(n_t, n_c)$-1] provides a p value, the probability of rejecting the null hypothesis. A p-value of ≤ 0.01, signifying a 99 percent probability the mean expression levels are different between the two groups (a 1% chance that the mean expression levels are in fact not different and that the observed difference occurred by statistical chance), is often considered acceptable.

When performing tests on a large scale, for example, on a large dataset of about 8000 genes, a correction factor must be included to adjust for the number of individual tests being performed. The most common and simplest correction is the

Bonferroni correction for multiple tests, which divides the p-value by the number of tests run. Using this test on an 8000 member dataset indicates that a p value of ≤0.00000125 is required to identify genes that are likely to be truly different between the two test conditions.

Wilcoxon's signed ranks test

This method is non-parametric and is utilized for paired comparisons. See e.g., Sokal and Rohlf (1987) Introduction to Biostatistics 2nd edition, WH Freeman, New York. At least 6 pairs are necessary to apply this statistic. This test is useful for analysis of paired expression data (for example, a set of patients who have cardiac transplant biopsy on 2 occasions and have a grade 0 on one occasion and a grade 3A on another).

ANOVA

Differences in gene expression across multiple related groups may be assessed using an Analysis of Variance (ANOVA), a method well known in the art (Michelson and Schofield, 1996).

Multivariate analysis

Many algorithms suitable for multivariate analysis are known in the art. Generally, a set of two or more genes for which expression discriminates between two disease states more specifically than expression of any single gene is identified by searching through the possible combinations of genes using a criterion for discrimination, for example the expression of gene X must increase from normal 300 percent, while the expression of genes Y and Z must decrease from normal by 75 percent. Ordinarily, the search starts with a single gene, then adds the next best fit at each step of the search. Alternatively, the search starts with all of the genes and genes that do not aid in the discrimination are eliminated step-wise.

Paired samples

Paired samples, or samples collected at different time-points from the same patient, are often useful, as described above. For example, use of paired samples permits the reduction of variation due to genetic variation among individuals. In addition, the use of paired samples has a statistical significance, in that data derived from paired samples can be calculated in a different manner that recognizes the reduced variability. For example, the formula for a t-test for paired samples is:

$$t(e_x) = \frac{\overline{D}_{e_x}}{\sqrt{\frac{\sum D^2 - (\sum D)^2 / b}{b-1}}}, \qquad (0.5)$$

where D is the difference between each set of paired samples and b is the number of sample pairs. $\overline{D}$ is the mean of the differences between the members of the pairs. In this test, only the differences between the paired samples are considered, then grouped together (as opposed to taking all possible differences between groups, as would be the case with an ordinary t-test). Additional statistical tests useful with paired data, e.g., ANOVA and Wilcoxon's signed rank test, are discussed above.

Diagnostic classification

Once a discriminating set of genes is identified, the diagnostic classifier (a mathematical function that assigns samples to diagnostic categories based on expression data) is applied to unknown sample expression levels.

Methods that can be used for this analysis include the following non-limiting list:

CLEAVER is an algorithm used for classification of useful expression profile data. See Raychaudhuri et al. (2001) Trends Biotechnol 19:189-193. CLEAVER uses positive training samples (e.g., expression profiles from samples known to be derived from a particular patient or sample diagnostic category, disease or disease criteria), negative training samples (e.g., expression profiles from samples known not to be derived from a particular patient or sample diagnostic category, disease or disease criteria) and test samples (e.g., expression profiles obtained from a patient), and determines whether the test sample correlates with the particular disease or disease criteria, or does not correlate with a particular disease or disease criteria. CLEAVER also generates a list of the 20 most predictive genes for classification.

Artificial neural networks (hereinafter, "ANN") can be used to recognize patterns in complex data sets and can discover expression criteria that classify samples into more than 2 groups. The use of artificial neural networks for discovery of gene expression diagnostics for cancers using expression data generated by oligonucleotide expression microarrays is demonstrated by Khan et al. (2001) Nature Med. 7:673-9. Khan found that 96 genes provided 0% error rate in classification of the tumors. The most important of these genes for classification was then determined

by measuring the sensitivity of the classification to a change in expression of each gene. Hierarchical clustering using the 96 genes results in correct grouping of the cancers into diagnostic categories.

Golub uses cDNA microarrays and a distinction calculation to identify genes with expression behavior that distinguishes myeloid and lymphoid leukemias. See Golub et al. (1999) Science 286:531-7. Self organizing maps were used for new class discovery. Cross validation was done with a "leave one out" analysis. 50 genes were identified as useful markers. This was reduced to as few as 10 genes with equivalent diagnostic accuracy.

Hierarchical and non-hierarchical clustering methods are also useful for identifying groups of genes that correlate with a subset of clinical samples such as with transplant rejection grade. Alizadeh used hierarchical clustering as the primary tool to distinguish different types of diffuse B-cell lymphomas based on gene expression profile data. See Alizadeh et al. (2000) Nature 403:503-11. Alizadeh used hierarchical clustering as the primary tool to distinguish different types of diffuse B-cell lymphomas based on gene expression profile data. A cDNA array carrying 17856 probes was used for these experiments, 96 samples were assessed on 128 arrays, and a set of 380 genes was identified as being useful for sample classification.

Perou demonstrates the use of hierarchical clustering for the molecular classification of breast tumor samples based on expression profile data. See Perou el al. (2000) Nature 406:747-52. In this work, a cDNA array carrying 8102 gene probes was used. 1753 of these genes were found to have high variation between breast tumors and were used for the analysis.

Hastie describes the use of gene shaving for discovery of expression markers. Hastie et al. (2000) Genome Biol. 1(2):RESEARCH 0003.1-0003.21. The gene shaving algorithm identifies sets of genes with similar or coherent expression patterns, but large variation across conditions (RNA samples, sample classes, patient classes). In this manner, genes with a tight expression pattern within a transplant rejection grade, but also with high variability across rejection grades are grouped together. The algorithm takes advantage of both characteristics in one grouping step. For example, gene shaving can identify useful marker genes with co-regulated expression. Sets of useful marker genes can be reduced to a smaller set, with each gene providing some non-redundant value in classification. This algorithm was used on the data set

described in Alizadeh et al., supra, and the set of 380 informative gene markers was reduced to 234.

Selected Diseases

In principle, diagnostic nucleotide sets of the invention may be developed and applied to essentially any disease, or disease criterion, as long as at least one subset of nucleotide sequences is differentially expressed in samples derived from one or more individuals with a disease criteria or disease and one or more individuals without the disease criteria or disease, wherein the individual may be the same individual sampled at different points in time, or the individuals may be different individuals (or populations of individuals). For example, the subset of nucleotide sequences may be differentially expressed in the sampled tissues of subjects with the disease or disease criterion (e.g., a patient with a disease or disease criteria) as compared to subjects without the disease or disease criterion (e.g., patients without a disease (control patients)). Alternatively, or in addition, the subset of nucleotide sequence(s) may be differentially expressed in different samples taken from the same patient, e.g at different points in time, at different disease stages, before and after a treatment, in the presence or absence of a risk factor, etc.

Expression profiles corresponding to sets of nucleotide sequences that correlate not with a diagnosis, but rather with a particular aspect of a disease can also be used to identify the diagnostic nucleotide sets and disease specific target nucleotide sequences of the invention. For example, such an aspect, or disease criterion, can relate to a subject's medical or family history, e.g., childhood illness, cause of death of a parent or other relative, prior surgery or other intervention, medications, symptoms (including onset and/or duration of symptoms), etc. Alternatively, the disease criterion can relate to a diagnosis, e.g., hypertension, diabetes, atherosclerosis, or prognosis (e.g., prediction of future diagnoses, events or complications), e.g., acute myocardial infarction, restenosis following angioplasty, reperfusion injury, allograft rejection, rheumatoid arthritis or systemic lupus erythematosis disease activity or the like. In other cases, the disease criterion corresponds to a therapeutic outcome, e.g., transplant rejection, bypass surgery or response to a medication, restenosis after stent implantation, collateral vessel growth due to therapeutic angiogenesis therapy, decreased angina due to revascularization, resolution of symptoms associated with a myriad of therapies, and the like. Alternatively, the disease criteria corresponds with

previously identified or classic risk factors and may correspond to prognosis or future disease diagnosis. As indicated above, a disease criterion can also correspond to genotype for one or more loci. Disease criteria (including patient data) may be collected (and compared) from the same patient at different points in time, from different patients, between patients with a disease (criterion) and patients respresenting a control population, etc. Longitudinal data, i.e., data collected at different time points from an individual (or group of individuals) may be used for comparisons of samples obtained from an individual (group of individuals) at different points in time, to permit identification of differences specifically related to the disease state, and to obtain information relating to the change in expression over time, including a rate of change or trajectory of expression over time. The usefulness of longitudinal data is further discussed in the section titled "Identification of diagnostic nucleotide sets of the invention".

It is further understood that diagnostic nucleotide sets may be developed for use in diagnosing conditions for which there is no present means of diagnosis. For example, in rheumatoid arthritis, joint destruction is often well under way before a patient experience symptoms of the condition. A diagnostic nucleotide set may be developed that diagnoses rheumatic joint destruction at an earlier stage than would be possible using present means of diagnosis, which rely in part on the presentation of symptoms by a patient. Diagnostic nucleotide sets may also be developed to replace or augment current diagnostic procedures. For example, the use of a diagnostic nucleotide set to diagnose cardiac allograft rejection may replace the current diagnostic test, a graft biopsy.

It is understood that the following discussion of diseases is exemplary and non-limiting, and further that the general criteria discussed above, e.g. use of family medical history, are generally applicable to the specific diseases discussed below.

In addition to leukocytes, as described throughout, the general method is applicable to nucleotide sequences that are differentially expressed in any subject tissue or cell type, by the collection and assessment of samples of that tissue or cell type. However, in many cases, collection of such samples presents significant technical or medical problems given the current state of the art.

Organ transplant rejection and success

A frequent complication of organ transplantation is recognition of the transplanted organ as foreign by the immune system resulting in rejection. Diagnostic

nucleotide sets can be identified and validated for monitoring organ transplant success, rejection and treatment. Medications currently exist that suppress the immune system, and thereby decrease the rate of and severity of rejection. However, these drugs also suppress the physiologic immune responses, leaving the patient susceptible to a wide variety of opportunistic infections. At present there is no easy, reliable way to diagnose transplant rejection. Organ biopsy is the preferred method, but this is expensive, painful and associated with significant risk and has inadequate sensitivity for focal rejection.

Diagnostic nucleotide sets of the present invention can be developed and validated for use as diagnostic tests for transplant rejection and success. It is appreciated that the methods of identifying diagnostic nucleotide sets are applicable to any organ transplant population. For example, diagnostic nucleotide sets are developed for cardiac allograft rejection and success. In some cases, disease criteria correspond to acute stage rejection diagnosis based on organ biopsy and graded using the International Society for Heart and Lung Transplantation ("ISHLT") criteria. Other disease criteria correspond to information from the patient's medical history and information regarding the organ donor. Alternatively, disease criteria include the presence or absence of cytomegalovirus (CMV) infection, Epstein-Barr virus (EBV) infection, allograft dysfunction measured by physiological tests of cardiac function (e.g., hemodynamic measurements from catheterization or echocardiograph data), and symptoms of other infections. Alternatively, disease criteria corresponds to therapeutic outcome, e.g. graft failure, re-transplantation, transplant vasculopathy, response to immunosuppressive medications, etc. Disease criteria may further correspond to a rejection episode of at least moderate histologic grade, which results in treatment of the patient with additional corticosteroids, anti-T cell antibodies, or total lymphoid irradiation; a rejection with histologic grade 2 or higher; a rejection with histologic grade <2; the absence of histologic rejection and normal or unchanged allograft function (based on hemodynamic measurements from catheterization or on echocardiographic data); the presence of severe allograft dysfunction or worsening allograft dysfunction during the study period (based on hemodynamic measurements from catheterization or on echocardiographic data).; documented CMV infection by culture, histology, or PCR, and at least one clinical sign or symptom of infection; specific graft biopsy rejection grades; rejection of mild to moderate histologic severity prompting augmentation of the patient's chronic immunosuppressive regimen;

rejection of mild to moderate severity with allograft dysfunction prompting plasmaphoresis or a diagnosis of "humoral" rejection; infections other than CMV, especially infection with Epstein Barr virus (EBV); lymphoproliferative disorder (also called post-transplant lymphoma); transplant vasculopathy diagnosed by increased intimal thickness on intravascular ultrasound (IVUS), angiography, or acute myocardial infarction; graft failure or retransplantation; and all cause mortality. Further specific examples of clinical data useful as disease criteria are provided in Example 11.

In another example, diagnostic nucleotide sets are developed and validated for use in treatment of kidney allograft rejection. Disease criteria correspond to, e.g., results of biopsy analysis for kidney allograft rejection, serum creatine level, and urinalysis results. Another disease criteria corresponds to the need for hemodialysis or other renal replacement therapy. Diagnostic nucleotide sets are developed and validated for use in diagnosis and treatment of bone marrow transplant rejection and liver transplant rejection, respectively. Disease criteria for bone marrow transplant rejection correspond to the diagnosis and monitoring of graft rejection and/or graft versus host disease. Disease criteria for liver transplant rejection include levels of serum markers for liver damage and liver function such as AST (aspartate aminotransferase), ALT (alanine aminotransferase), Alkaline phosphatase, GGT, (gamma-glutamyl transpeptidase) Bilirubin, Albumin and Prothrombin time. Further disease criteria correspond to hepatic encephalopathy, medication usage, ascites, and histological rejection on graft biopsy. In addition, urine can be utilized for at the target tissue for profiling in renal transplant, while biliary and intestinal and feces may be used favorably for hepatic or intestinal organ allograft rejection.

Atherosclerosis and Stable Angina Pectoris

Over 50 million patients in the U.S. have atherosclerotic coronary artery disease (hereinafter, "CAD"), and it is of great importance to identify patients who will suffer complications from the disease. Atherosclerosis leads to progressive narrowing of the coronary arteries, which may lead to myocardial ischemia, which manifests as stable angina pectoris, or chest pain with exertion. In addition to chest pain, patients may also have shortness of breath (dyspnea), fatigue, nausea or other symptoms with exertion. Myocardial infarction (heart attack) and unstable angina are acute events associated with atherosclerosis. There is currently no way to accurately predict the occurrence of acute events in patients with atherosclerosis, however.

Although the presence of classic risk factors and arterial wall calcification (as assessed by CT scanning) is weakly correlated with the occurrence of acute coronary syndrome, the degree of artery stenosis (i.e. vessel occlusion as a result of atherosclerosis) correlates poorly with the occurrence of future acute events, as acute events occur more commonly in coronary arteries with 40-50% blockage than arteries that are 80-90% blocked. Coronary angiography can provide information about degree of coronary blockage, but is a poor tool for the measurement of disease activity and the prediction of the likelihood of acute events and other poor outcomes.

Diagnostic nucleotide sets are developed and validated for use in diagnosis and monitoring of atherosclerosis, and in predicting the likelihood of complications, e.g. angina and myocardial infarction. Alternatively, or in addition, disease criteria correspond to symptoms or diagnosis of disease progression, e.g. clinical results of angiography indicating progressive narrowing of vessel lumens. In another aspect, diagnostic nucleotide sets are developed for use in predicting the likelihood of future acute events in patients suffering from atherosclerosis. Disease criteria correspond to retrospective data, for example a recent history of unstable angina or myocardial infarction. Disease criteria also correspond to prospective data, for example, the occurrence of unstable angina or myocardial infarction. In another case, disease criteria correspond to standard medical indicators of occurrence of an acute event, e.g. serum enzyme levels, electrocardiographic testing, chest pain, nuclear magnetic imaging, etc.

Congestive Heart Failure

Congestive heart failure (hereinafter, "CHF") is a disease that affects increasing numbers of individuals. Without being bound by theory, it is believed that CHF is associated with systemic inflammation. Markers of systemic inflammation and serum cytokine levels such as erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) and serum cytokine levels are elevated (or altered) in patients with CHF, and elevation correlates with the severity and progression of the disease. Furthermore, serum catecholamine levels (epinephrine and norepinephrine) are also elevated in proportion to the severity of CHF, and may directly alter leukocyte expression patterns. Currently, echocardiography is the test primarily used to assess the severity of CHF and monitor progression of the disease. There are a number of drugs that are efficacious in treating CHF, such as beta-blockers and ACE inhibitors.

A leukocyte test with the ability to determine the rate of progression and the adequacy of therapy is of great interest.

Diagnostic nucleotide sets are developed and validated for use in diagnosis and monitoring of progression and rate of progression (activity) of CHF. Disease criteria correspond to the results of echocardiography testing, which may indicate diagnosis of CHF or increasing severity of CHF as evidenced by worsening parameters for ventricular function, such as the ejection fraction, fractional shortening, wall motion or ventricular pressures. Alternatively, or in addition, disease criteria correspond to hospitalization for CHF, death, pulmonary edema, increased cardiac chamber dimensions on echocardiography or another imaging test, exercise testing of hemodynamic measurements, serial CRP, other serum markers, NYHA functional classes, quality of life measures, renal function, transplant listing, pulmonary edema, left ventricular assist device use, medication use and changes, and worsening of Ejection Fraction by echocardiography, angiography, MRI, CT or nuclear imaging.. In another aspect, disease criteria correspond to response to drug therapy, e.g. beta-blockers or ACE inhibitors.

Risk factors for coronary artery disease

The established and classic risks for the occurrence of coronary artery disease and complications of that disease are: cigarette smoking, diabetes, hypertension, hyperlipidemia and a family history of early atherosclerosis. Obesity, sedentary lifestyle, syndrome X, cocaine use, chronic hemodialysis and renal disease, radiation exposure, endothelial dysfunction, elevated plasma homocysteine, elevated plasma lipoprotein a, elevated CRP, infection with CMV and chlamydia infection are less well established, controversial, or putative risk factors for the disease. Risk factors are known to be associated with patient prognosis and outcome, but the contribution of each risk factor to the future clinical state of a patient is difficult to measure. The effect of risk factor modification (e.g., smoking cessation, treatment of hypercholesterolemia) on overall risk and future outcome is also difficult to quantify.

Diagnostic nucleotide sets may be developed that correlate with these risk factors, or the sum of the risk factors for use in predicting occurrence of coronary artery disease. Disease criteria correspond to risk factors, as exemplified above, as well as to occurrence of coronary artery disease. Alternatively, or in addition, disease criteria corresponding to risk factors may contribute to a numerical weighted average, which itself may be treated as a disease criteria and may be used for correlation to

gene expression. In another aspect, risk factors may be modified in a patient, e.g. by behavioral change, or decrease cholesterol through chemotherapy in patients with hypocholesteremia. Disease criteria may further correspond to diagnosis of coronary disease.

Restenosis

Angioplasty can re-open a narrowed artery. However, the long-term success rate of these procedures is limited by restenosis, the re-narrowing of a coronary artery after an angioplasty. Currently, about 50% of treated arteries re-narrow after angioplasty and about 30% re-narrow after standard stent placement. Restenosis usually becomes apparent within 3 months of the angioplasty procedure. Presently, there is no reliable method for predicting which arteries will succumb to restenosis, though small vessels tend to be more likely to re-narrow, as do vessels of diabetics, renal patients and vessels exposed to high-pressure balloon inflation during balloon angioplasty.

Diagnostic nucleotide sets are developed and validated to predict restenosis in patients before undergoing angioplasty or shortly thereafter. Disease criteria correspond to angiogram testing (diagnosis of restenosis) , as well as clinical symptoms of restenosis, e.g. chest pain due to re-narrowing of the artery, as confirmed by angiogram. Anti-restenotic drug therapy is also identified for each patient. The diagnostic nucleotide set are useful to identify patients about to undergo angioplasty who would benefit from stents, radiation-emitting stents, and anti-restenotic drug delivering stents. Patients that would benefit from post-angioplasty anti-restenotic drug therapy may also be identified.

Rheumatoid Arthritis

Rheumatoid arthritis (RA) effects about two million patients in the US and is a chronic and debilitating inflammatory arthritis, particularly involving pain and destruction of the joints. RA often goes undiagnosed because patients may have no pain, but the disease is actively destroying the joint. Other patients are known to have RA, and are treated to alleviate symptoms, but the rate of progression of joint destruction can't easily be monitored. Drug therapy is available, but the most effective medicines are toxic (e.g., steroids, methotrexate) and thus need to be used with caution. A new class of medications (TNF blockers) is very effective, but the drugs are expensive, have side effects, and not all patients respond. Side-effects are

common and include immune suppression, toxicity to organ systems, allergy and metabolic disturbances.

Diagnostic nucleotide sets of the invention are developed and validated for use in diagnosis and treatment of RA. Disease criteria correspond to disease symptoms (e.g., joint pain, joint swelling and joint stiffness and any of the American College for Rheumatology criteria for the diagnosis of RA, see Arnett et al (1988) Arthr. Rheum. 31:315-24), progression of joint destruction (e.g. as measured by serial hand radiographs, assessment of joint function and mobility), surgery, need for medication, additional diagnoses of inflammatory and non-inflammatory conditions, and clinical laboratory measurements including complete blood counts with differentials, CRP, ESR, ANA, Serum IL6, Soluble CD40 ligand, LDL, HDL, Anti-DNA antibodies, rheumatoid factor, C3, C4, serum creatinine. In addition, or alternatively, disease criteria correspond to response to drug therapy and presence or absence of side-effects or measures of improvement exemplified by the American College of Rheumatology "20%" and "50%" response/improvement rates. See Felson et al (1995) Arthr Rheum 38:531-37. Diagnostic nucleotide sets are identified that monitor and predict disease progression including flaring (acute worsening of disease accompanied by joint pain or other symptoms), response to drug treatment and likelihood of side-effects.

In addition to peripheral leukocytes, surgical specimens of rheumatoid joints can be used for leukocyte expression profiling experiments. Members of diagnostic nucleotide sets are candidates for leukocyte target nucleotide sequences, e.g. as a candidate drug target for rheumatoid arthritis.

Systemic Lupus Erythematosis (SLE)

SLE is a chronic, systemic inflammatory disease characterized by dysregulation of the immune system, which effects up to 2 million patients in the US. Symptoms of SLE include rashes, joint pain, abnormal blood counts, renal dysfunction and damage, infections, CNS disorders, arthralgias and autoimmunity. Patients may also have early onset atherosclerosis.

Diagnostic nucleotide sets are identified and validated for use in diagnosis and monitoring of SLE activity and progression. Disease criteria correspond to clinical data, e.g. symptom rash, joint pain, malaise, rashes, blood counts (white and red), tests of renal function e.g. creatinine, blood urea nitrogen (hereinafter, "bun") creative clearance, data obtained from laboratory tests including complete blood counts with differentials, CRP, ESR, ANA, Serum IL6, Soluble CD40 ligand, LDL, HDL, Anti-

DNA antibodies, rheumatoid factor, C3, C4, serum creatinine and any medication levels, the need for pain medications, cumulative doses or immunosuppressive therapy, symptoms or any manifestation of carotid atherosclerosis (e.g. ultrasound diagnosis or any other manifestations of the disease), data from surgical procedures such as gross operative findings and pathological evaluation of resected tissues and biopsies (e.g., renal, CNS), information on pharmacological therapy and treatment changes, clinical diagnoses of disease "flare", hospitalizations, death, quantitative joint exams, results from health assessment questionnaires (HAQs), and other clinical measures of patient symptoms and disability. In addition, disease criteria correspond to the clinical score known as SLEDAI (Bombadier C, Gladman DD, Urowitz MB, Caron D, Chang CH and the Committee on Prognosis Studies in SLE: Derivation of the SLEDAI for Lupus Patients. Arthritis Rheum 35:630-640, 1992.). Diagnostic nucleotide sets may be useful for diagnosis of SLE, monitoring disease progression including progressive renal dysfunction, carotid atherosclerosis and CNS dysfunction, and predicting occurrence of side-effects, for example.

Dermatomyositis/Polymyositis

Dermatomyositis/Polymyositis is an autoimmune/inflammatory disease of muscle and skin. Disease criteria correspond to clinical markers of muscle damage (e.g. creatine kinase or myoglobin), muscle strength, symptoms, skin rash or muscle biopsy results.

Diabetes

Insulin dependent (type I) diabetes is caused by an autoimmune attack of insulin producing cells in the pancreas. The disease does not manifest until greater than 90% of the insulin producing cells are destroyed. Diagnostic nucleotide sets are developed and validated for use in detecting diabetes before it is clinically evident. Disease criteria correspond to future occurrence of diabetes, glucose tolerance, serum glucose level, and levels of hemoglobin A1c or other markers.

Inflammatory Bowel Disease (Crohn's and Ulcerative Colitis)

Inflammatory Bowel Disease, e.g., Crohn's Disease and Ulcerative Colitis, are chronic inflammatory diseases of the intestine. Together they effect at least 1 million in the US. Currently, diagnosis and monitoring is accomplished by intestinal endoscopy with or without a biopsy. Steroids and other immune suppressing drugs are useful in treating these diseases, but these drugs cause toxicity and severe side-effects. Diagnostic nucleotide sets are developed for use in diagnosis and monitoring

of disease progression. Disease criteria correspond to clinical criteria, e.g. symptoms of abdominal or pelvic pain, diarrhea, fever and rectal bleeding. Alternatively, or in addition, disease criteria correspond to endoscopy results or bowel biopsy results.

Osteoarthritis

20-40 million patients in the US have osteoarthritis. Patient groups are heterogeneous, with a subset of patients having earlier onset, more aggressive joint damage, involving more inflammation (leukocyte infiltration) leukocyte diagnostics can be used to distinguish osteoarthritis from rheumatoid arthritis, define likelihood and degree of response to NSAID therapy (non-steroidal anti-inflammatory drugs). Rate of progression of joint damage can also be assessed. Diagnostic nucleotide sets may be developed for use in selection and titration of treatment therapies. Disease criteria correspond to response to therapy, and disease progression using certain therapies, need for joint surgery, joint pain and disability.

Asthma

Asthma is a chronic inflammatory disease of the lungs. Clinical symptoms include chronic or acute airflow obstruction. Patients are treated with inhaled steroids or bronchodilators or systemic steroids and other medication, and disease progression is monitored clinically using a peak air flow meter or formal pulmonary function tests. Even with these tests, it is difficult to predict which patients are at highest risk for acute worsening of airway obstruction (an "asthma attack"). Diagnostic nucleotide sets are developed for use in predicting likelihood of acute asthma attacks, and for use in choosing and titrating drug therapy. Disease criteria correspond to pulmonary function testing, peak flow meter measurements, ER visits, inhaler use, subjective patient assessment of response to therapy, hospitalization and need for steroids.

Other inflammatory diseases:

Other inflammatory disease suitable for development and use of diagnostic nucleotide sets are polymyalgia rheumatica, temporal arteritis, polyarteritis nodosa, wegener's granulomatosis, whipple's disease, heterotopic ossification, Periprosthetic Osteolysis, Sepsis/ARDS, scleroderma, Grave's disease, Hashimoto's thyroiditis, psoriasis numerous others (See Table 1).

Viral diseases

Diagnostic leukocyte nucleotide sets may be developed and validated for use in diagnosing viral disease. In another aspect, viral nucleotide sequences may be

added to a leukocyte nucleotide set for use in diagnosis of viral diseases. Alternatively, viral nucleotide sets and leukocyte nucleotides sets may be used sequentially.

Epstein-Barr virus (EBV)

EBV causes a variety of diseases such as mononucleosis, B-cell lymphoma, and pharyngeal carcinoma. It infects mononuclear cells and circulating atypical lymphocytes are a common manifestation of infection. Peripheral leukocyte gene expression is altered by infection. Transplant recipients and patients who are immunosuppressed are at increased risk for EBV-associated lymphoma.

Diagnostic nucleotide sets may be developed and validated for use in diagnosis and monitoring of EBV. In one aspect, the diagnostic nucleotide set is a leukocyte nucleotide set. Alternatively, EBV nucleotide sequences are added to a leukocyte nucleotide set, for use in diagnosing EBV. Disease criteria correspond with diagnosis of EBV, and, in patients who are EBV-sero-positive, presence (or prospective occurrence) of EBV-related illnesses such as mononucleosis, and EBV-associated lymphoma. Diagnostic nucleotide sets are useful for diagnosis of EBV, and prediction of occurrence of EBV-related illnesses.

Cytomegalovirus (CMV)

Cytomegalovirus cause inflammation and disease in almost any tissue, particularly the colon, lung, bone marrow and retina, and is a very important cause of disease in immunosuppressed patients, e.g. transplant, cancer, AIDS. Many patients are infected with or have been exposed to CMV, but not all patients develop clinical disease from the virus. Also, CMV negative recipients of allografts that come from CMV positive donors are at high risk for CMV infection. As immunosuppressive drugs are developed and used, it is increasingly important to identify patients with current or impending clinical CMV disease, because the potential benefit of immunosuppressive therapy must be balanced with the increased rate of clinical CMV infection and disease that may result from the use of immunosuppression therapy. CMV may also play a role in the occurrence of atherosclerosis or restenosis after angioplasty.

Diagnostic nucleotide sets are developed for use in diagnosis and monitoring of CMV infection or re-activation of CMV infection. In one aspect, the diagnostic nucleotide set is a leukocyte nucleotide set. In another aspect, CMV nucleotide sequences are added to a leukocyte nucleotide set, for use in diagnosing CMV.

Disease criteria correspond to diagnosis of CMV (e.g., sero-positive state) and presence of clinically active CMV. Disease criteria may also correspond to prospective data, e.g. the likelihood that CMV will become clinically active or impending clinical CMV infection. Antiviral medications are available and diagnostic nucleotide sets can be used to select patients for early treatment, chronic suppression or prophylaxis of CMV activity.

Hepatitis B and C

These chronic viral infections affect about 1.25 and 2.7 million patients in the US, respectively. Many patients are infected, but suffer no clinical manifestations. Some patients with infection go on to suffer from chronic liver failure, cirrhosis and hepatic carcinoma.

Diagnostic nucleotide sets are developed for use in diagnosis and monitoring of HBV or HCV infection. In one aspect, the diagnostic nucleotide set is a leukocyte nucleotide set. In another aspect, viral nucleotide sequences are added to a leukocyte nucleotide set, for use in diagnosing the virus and monitoring progression of liver disease. Disease criteria correspond to diagnosis of the virus (e.g., sero-positive state or other disease symptoms). Alternatively, disease criteria correspond to liver damage, e.g., elevated alkaline phosphatase, ALT, AST or evidence of ongoing hepatic damage on liver biopsy. Alternatively, disease criteria correspond to serum liver tests (AST, ALT, Alkaline Phosphatase, GGT, PT, bilirubin), liver biopsy, liver ultrasound, viral load by serum PCR, cirrhosis, hepatic cancer, need for hospitalization or listing for liver transplant. Diagnostic nucleotide sets are used to diagnose HBV and HCV, and to predict likelihood of disease progression. Antiviral therapeutic usage, such as Interferon gamma and Ribavirin, can also be disease criteria.

HIV

HIV infects T cells and certainly causes alterations in leukocyte expression. Diagnostic nucleotide sets are developed for diagnosis and monitoring of HIV. In one aspect, the diagnostic nucleotide set is a leukocyte nucleotide set. In another aspect, viral nucleotide sequences are added to a leukocyte nucleotide set, for use in diagnosing the virus. Disease criteria correspond to diagnosis of the virus (e.g., sero-positive state). In addition, disease criteria correspond to viral load, CD4 T cell counts, opportunistic infection, response to antiretroviral therapy, progression to AIDS, rate of progression and the occurrence of other HIV related outcomes (e.g.,

malignancy, CNS disturbance). Response to antiretrovirals may also be disease criteria.

Pharmacogenomics

Pharmocogenomics is the study of the individual propensity to respond to a particular drug therapy (combination of therapies). In this context, response can mean whether a particular drug will work on a particular patient, e.g. some patients respond to one drug but not to another drug. Response can also refer to the likelihood of successful treatment or the assessment of progress in treatment. Titration of drug therapy to a particular patient is also included in this description, e.g. different patients can respond to different doses of a given medication. This aspect may be important when drugs with side-effects or interactions with other drug therapies are contemplated.

Diagnostic nucleotide sets are developed and validated for use in assessing whether a patient will respond to a particular therapy and/or monitoring response of a patient to drug therapy(therapies). Disease criteria correspond to presence or absence of clinical symptoms or clinical endpoints, presence of side-effects or interaction with other drug(s). The diagnostic nucleotide set may further comprise nucleotide sequences that are targets of drug treatment or markers of active disease.

Validation and accuracy of diagnostic nucleotide set using correlation analysis

Prior to widespread application of the diagnostic probe sets of the invention, the predictive value of the probe set is validated.

Typically, the oligonucleotide sequence of each probe is confirmed, e.g. by DNA sequencing using an oligonucleotide-specific primer. Partial sequence obtained is generally sufficient to confirm the identity of the oligonucleotide probe. Alternatively, a complementary polynucleotide is fluorescently labeled and hybridized to the array, or to a different array containing a resynthesized version of the oligo nucleotide probe, and detection of the correct probe is confirmed.

Typically, validation is performed by statistically evaluating the accuracy of the correspondence between the molecular signature for a diagnostic probe set and a selected indicator. For example, the expression differential for a nucleotide sequence between two subject classes can be expressed as a simple ratio of relative expression. The expression of the nucleotide sequence in subjects with selected indicator can be

compared to the expression of that nucleotide sequence in subjects without the indicator, as described in the following equations.

$\sum E_x ai/N = E_x A$ the average expression of nucleotide sequence x in the members of group A;

$\sum E_x bi/M = E_x B$ the average expression of nucleotide sequence x in the members of group B;

$E_x A/ ExB = \Delta E_x AB$ the average differential expression of nucleotide sequence x between groups A

and B:

where $\sum$ indicates a sum; Ex is the expression of nucleotide sequence x relative to a standard; ai are the individual members of group A, group A has N members; bi are the individual members of group B, group B has M members.

The expression of at least two nucleotide sequences, e.g., nucleotide sequence X and nucleotide sequence Y are measured relative to a standard in at least one subject of group A (e.g., with a disease) and group B (e.g., without the disease). Ideally, for purposes of validation the indicator is independent from (i.e., not assigned based upon) the expression pattern. Alternatively, a minimum threshold of gene expression for nucleotide sequences X and Y, relative to the standard, are designated for assignment to group A. For nucleotide sequence x, this threshold is designated ΔEx, and for nucleotide sequence y, the threshold is designated ΔEy.

The following formulas are used in the calculations below:

Sensitivity = (true positives/true positives + false negatives)

Specificity = (true negatives/true negatives + false positives)

If, for example, expression of nucleotide sequence x above a threshold: $x > \Delta Ex$, is observed for 80/100 subjects in group A and for 10/100 subjects in group B, the sensitivity of nucleotide sequence x for the assignment to group A, at the given expression threshold ΔEx, is 80%, and the specificity is 90%.

If the expression of nucleotide sequence y is $> \Delta Ey$ in 80/100 subjects in group A, and in 10/100 subjects in group B, then, similarly the sensitivity of nucleotide sequence y for the assignment to group A at the given threshold ΔEy is 80% and the specificity is 90%. If in addition, 60 of the 80 subjects in group A that meet the expression threshold for nucleotide sequence y also meet the expression threshold ΔEx and that 5 of the 10 subjects in group B that meet the expression

threshold for nucleotide sequence y also meet the expression threshold ΔEx, the sensitivity of the test (x>ΔEx and y>ΔEy)for assignment of subjects to group A is 60% and the specificity is 95%.

Alternatively, if the criteria for assignment to group A are change to: Expression of x > ΔEx or expression of y > ΔEy, the sensitivity approaches 100% and the specificity is 85%.

Clearly, the predictive accuracy of any diagnostic probe set is dependent on the minimum expression threshold selected. The expression of nucleotide sequence X (relative to a standard) is measured in subjects of groups A (with disease) and B (without disease). The minimum threshold of nucleotide sequence expression for x, required for assignment to group A is designated ΔEx 1.

If 90/100 patients in group A have expression of nucleotide sequence x > ΔEx 1 and 20/100 patients in group B have expression of nucleotide sequence x > ΔEx 1, then the sensitivity of the expression of nucleotide sequence x (using ΔEx 1 as a minimum expression threshold) for assignment of patients to group A will be 90% and the specificity will be 80%.

Altering the minimum expression threshold results in an alteration in the specificity and sensitivity of the nucleotide sequences in question. For example, if the minimum expression threshold of nucleotide sequence x for assignment of subjects to group A is lowered to ΔEx 2, such that 100/100 subjects in group A and 40/100 subjects in group B meet the threshold, then the sensitivity of the test for assignment of subjects to group A will be 100% and the specificity will be 60%.

Thus, for 2 nucleotide sequences X and Y: the expression of nucleotide sequence x and nucleotide sequence y (relative to a standard) are measured in subjects belonging to groups A (with disease) and B (without disease). Minimum thresholds of nucleotide sequence expression for nucleotide sequences X and Y (relative to common standards) are designated for assignment to group A. For nucleotide sequence x, this threshold is designated ΔEx1 and for nucleotide sequence y, this threshold is designated ΔEy1.

If in group A, 90/100 patients meet the minimum requirements of expression ΔEx1 and ΔEy1, and in group B, 10/100 subjects meet the minimum requirements of expression ΔEx1 and ΔEy1, then the sensitivity of the test for assignment of subjects to group A is 90% and the specificity is 90%.

Increasing the minimum expression thresholds for X and Y to ΔEx2 and ΔEy2, such that in group A, 70/100 subjects meet the minimum requirements of expression ΔEx2 and ΔEy2, and in group B, 3/100 subjects meet the minimum requirements of expression ΔEx2 and ΔEy2. Now the sensitivity of the test for assignment of subjects to group A is 70% and the specificity is 97%.

If the criteria for assignment to group A is that the subject in question meets either threshold, ΔEx2 or ΔEy2, and it is found that 100/100 subjects in group A meet the criteria and 20/100 subjects in group B meet the criteria, then the sensitivity of the test for assignment to group A is 100% and the specificity is 80%.

Individual components of a diagnostic probe set each have a defined sensitivity and specificity for distinguishing between subject groups. Such individual nucleotide sequences can be employed in concert as a diagnostic probe set to increase the sensitivity and specificity of the evaluation. The database of molecular signatures is queried by algorithms to identify the set of nucleotide sequences (i.e., corresponding to members of the probe set) with the highest average differential expression between subject groups. Typically, as the number of nucleotide sequences in the diagnostic probe set increases, so does the predictive value, that is, the sensitivity and specificity of the probe set. When the probe sets are defined they may be used for diagnosis and patient monitoring as discussed below. The diagnostic sensitivity and specificity of the probe sets for the defined use can be determined for a given probe set with specified expression levels as demonstrated above. By altering the expression threshold required for the use of each nucleotide sequence as a diagnostic, the sensitivity and specificity of the probe set can be altered by the practitioner. For example, by lowering the magnitude of the expression differential threshold for each nucleotide sequence in the set, the sensitivity of the test will increase, but the specificity will decrease. As is apparent from the foregoing discussion, sensitivity and specificity are inversely related and the predictive accuracy of the probe set is continuous and dependent on the expression threshold set for each nucleotide sequence. Although sensitivity and specificity tend to have an inverse relationship when expression thresholds are altered, both parameters can be increased as nucleotide sequences with predictive value are added to the diagnostic nucleotide set. In addition a single or a few markers may not be reliable expression markers across a population of patients. This is because of the variability in expression and measurement of expression that exists between measurements, individuals and

individuals over time. Inclusion of a large number of candidate nucleotide sequences or large numbers of nucleotide sequences in a diagnostic nucleotide set allows for this variability as not all nucleotide sequences need to meet a threshold for diagnosis. Generally, more markers are better than a single marker. If many markers are used to make a diagnosis, the likelihood that all expression markers will not meet some thresholds based upon random variability is low and thus the test will give fewer false negatives.

It is appreciated that the desired diagnostic sensitivity and specificity of the diagnostic nucleotide set may vary depending on the intended use of the set. For example, in certain uses, high specificity and high sensitivity are desired. For example, a diagnostic nucleotide set for predicting which patient population may experience side effects may require high sensitivity so as to avoid treating such patients. In other settings, high sensitivity is desired, while reduced specificity may be tolerated. For example, in the case of a beneficial treatment with few side effects, it may be important to identify as many patients as possible (high sensitivity) who will respond to the drug, and treatment of some patients who will not respond is tolerated. In other settings, high specificity is desired and reduced sensitivity may be tolerated. For example, when identifying patients for an early-phase clinical trial, it is important to identify patients who may respond to the particular treatment. Lower sensitivity is tolerated in this setting as it merely results in reduced patients who enroll in the study or requires that more patients are screened for enrollment.

Methods of using diagnostic nucleotide sets.

The invention also provide methods of using the diagnostic nucleotide sets to: diagnose disease; assess severity of disease; predict future occurrence of disease; predict future complications of disease; determine disease prognosis; evaluate the patient's risk, or "stratify" a group of patients; assess response to current drug therapy; assess response to current non-pharmacological therapy; determine the most appropriate medication or treatment for the patient; predict whether a patient is likely to respond to a particular drug; and determine most appropriate additional diagnostic testing for the patient, among other clinically and epidemiologically relevant applications.

The nucleotide sets of the invention can be utilized for a variety of purposes by physicians, healthcare workers, hospitals, laboratories, patients, companies and

other institutions. As indicated previously, essentially any disease, condition, or status for which at least one nucleotide sequence is differentially expressed in leukocyte populations (or sub-populations) can be evaluated, e.g., diagnosed, monitored, etc. using the diagnostic nucleotide sets and methods of the invention. In addition to assessing health status at an individual level, the diagnostic nucleotide sets of the present invention are suitable for evaluating subjects at a "population level," e.g., for epidemiological studies, or for population screening for a condition or disease.

Collection and preparation of sample

RNA, protein and/or DNA is prepared using methods well-known in the art, as further described herein. It is appreciated that subject samples collected for use in the methods of the invention are generally collected in a clinical setting, where delays may be introduced before RNA samples are prepared from the subject samples of whole blood, e.g. the blood sample may not be promptly delivered to the clinical lab for further processing. Further delay may be introduced in the clinical lab setting where multiple samples are generally being processed at any given time. For this reason, methods which feature lengthy incubations of intact leukocytes at room temperature are not preferred, because the expression profile of the leukocytes may change during this extended time period. For example, RNA can be isolated from whole blood using a phenol/guanidine isothiocyanate reagent or another direct whole-blood lysis method, as described in, e.g., U.S. Patent Nos. 5,346,994 and 4,843,155. This method may be less preferred under certain circumstances because the large majority of the RNA recovered from whole blood RNA extraction comes from erythrocytes since these cells outnumber leukocytes 1000:1. Care must be taken to ensure that the presence of erythrocyte RNA and protein does not introduce bias in the RNA expression profile data or lead to inadequate sensitivity or specificity of probes.

Alternatively, intact leukocytes may be collected from whole blood using a lysis buffer that selectively lyses erythrocytes, but not leukocytes, as described, e.g., in (U.S. Patent Nos. 5,973,137, and 6,020,186). Intact leukocytes are then collected by centrifugation, and leukocyte RNA is isolated using standard protocols, as described herein. However, this method does not allow isolation of sub-populations of leukocytes, e.g. mononuclear cells, which may be desired. In addition, the expression profile may change during the lengthy incubation in lysis buffer, especially

in a busy clinical lab where large numbers of samples are being prepared at any given time.

Alternatively, specific leukocyte cell types can be separated using density gradient reagents (Boyum, A, 1968.). For example, mononuclear cells may be separated from whole blood using density gradient centrifugation, as described, e.g., in U.S. Patents Nos. 4190535, 4350593, 4751001, 4818418, and 5053134. Blood is drawn directly into a tube containing an anticoagulant and a density reagent (such as Ficoll or Percoll). Centrifugation of this tube results in separation of blood into an erythrocyte and granulocyte layer, a mononuclear cell suspension, and a plasma layer. The mononuclear cell layer is easily removed and the cells can be collected by centrifugation, lysed, and frozen. Frozen samples are stable until RNA can be isolated. Density centrifugation, however, must be conducted at room temperature, and if processing is unduly lengthy, such as in a busy clinical lab, the expression profile may change.

The quality and quantity of each clinical RNA sample is desirably checked before amplification and labeling for array hybridization, using methods known in the art. For example, one microliter of each sample may be analyzed on a Bioanalyzer (Agilent 2100 Palo Alto, CA. USA) using an RNA 6000 nano LabChip (Caliper, Mountain View, CA. USA). Degraded RNA is identified by the reduction of the 28S to 18S ribosomal RNA ratio and/or the presence of large quantities of RNA in the 25-100 nucleotide range.

It is appreciated that the RNA sample for use with a diagnostic nucleotide set may be produced from the same or a different cell population, sub-population and/or cell type as used to identify the diagnostic nucleotide set. For example, a diagnostic nucleotide set identified using RNA extracted from mononuclear cells may be suitable for analysis of RNA extracted from whole blood or mononuclear cells, depending on the particular characteristics of the members of the diagnostic nucleotide set. Generally, diagnostic nucleotide sets must be tested and validated when used with RNA derived from a different cell population, sub-population or cell type than that used when obtaining the diagnostic gene set. Factors such as the cell-specific gene expression of diagnostic nucleotide set members, redundancy of the information provided by members of the diagnostic nucleotide set, expression level of the member of the diagnostic nucleotide set, and cell-specific alteration of expression of a member of the diagnostic nucleotide set will contribute to the usefullness of using a different

RNA source than that used when identifying the members of the diagnostic nucleotide set. It is appreciated that it may be desirable to assay RNA derived from whole blood, obviating the need to isolate particular cell types from the blood.

Rapid method of RNA extraction suitable for production in a clinical setting of high quality RNA for expression profiling

In a clinical setting, obtaining high quality RNA preparations suitable for expression profiling, from a desired population of leukocytes poses certain technical challenges, including: the lack of capacity for rapid, high-throughput sample processing in the clinical setting, and the possibility that delay in processing (in a busy lab or in the clinical setting) may adversely affect RNA quality, e.g. by a permitting the expression profile of certain nucleotide sequences to shift. Also, use of toxic and expensive reagents, such as phenol, may be disfavored in the clinical setting due to the added expense associated with shipping and handling such reagents.

A useful method for RNA isolation for leukocyte expression profiling would allow the isolation of monocyte and lymphocyte RNA in a timely manner, while preserving the expression profiles of the cells, and allowing inexpensive production of reproducible high-quality RNA samples. Accordingly, the invention provides a method of adding inhibitor(s) of RNA transcription and/or inhibitor(s) of protein synthesis, such that the expression profile is "frozen" and RNA degradation is reduced. A desired leukocyte population or sub-population is then isolated, and the sample may be frozen or lysed before further processing to extract the RNA. Blood is drawn from subject population and exposed to ActinomycinD (to a final concentration of 10 ug/ml) to inhibit transcription, and cycloheximide (to a final concentration of 10 ug/ml) to inhibit protein synthesis. The inhibitor(s) can be injected into the blood collection tube in liquid form as soon as the blood is drawn, or the tube can be manufactured to contain either lyophilized inhibitors or inhibitors that are in solution with the anticoagulant. At this point, the blood sample can be stored at room temperature until the desired leukocyte population or sub-population is isolated, as described elsewhere. RNA is isolated using standard methods, e.g., as described above, or a cell pellet or extract can be frozen until further processing of RNA is convenient.

The invention also provides a method of using a low-temperature density gradient for separation of a desired leukocyte sample. In another embodiment, the invention provides the combination of use of a low-temperature density gradient and the use of transcriptional and/or protein synthesis inhibitor(s). A desired leukocyte population is separated using a density gradient solution for cell separation that maintains the required density and viscosity for cell separation at 0-4°C. Blood is drawn into a tube containing this solution and may be refrigerated before and during processing as the low temperatures slow cellular processes and minimize expression profile changes. Leukocytes are separated, and RNA is isolated using standard methods. Alternately, a cell pellet or extract is frozen until further processing of RNA is convenient. Care must be taken to avoid rewarming the sample during further processing steps.

Alternatively, the invention provides a method of using low-temperature density gradient separation, combined with the use of actinomycin A and cyclohexamide, as described above.

Assessing expression for diagnostics

Expression profiles for the set of diagnostic nucleotide sequences in a subject sample can be evaluated by any technique that determines the expression of each component nucleotide sequence. Methods suitable for expression analysis are known in the art, and numerous examples are discussed in the Sections titled "Methods of obtaining expression data" and "high throughput expression Assays", above.

In many cases, evaluation of expression profiles is most efficiently, and cost effectively, performed by analyzing RNA expression. Alternatively, the proteins encoded by each component of the diagnostic nucleotide set are detected for diagnostic purposes by any technique capable of determining protein expression, e.g., as described above. Expression profiles can be assessed in subject leukocyte sample using the same or different techniques as those used to identify and validate the diagnostic nucleotide set. For example, a diagnostic nucleotide set identified as a subset of sequences on a cDNA microarray can be utilized for diagnostic (or prognostic, or monitoring, etc.) purposes on the same array from which they were identified. Alternatively, the diagnostic nucleotide sets for a given disease or condition can be organized onto a dedicated sub-array for the indicated purpose. It is important to note that if diagnostic nucleotide sets are discovered using one

technology, e.g. RNA expression profiling, but applied as a diagnostic using another technology, e.g. protein expression profiling, the nucleotide sets must generally be validated for diagnostic purposes with the new technology. In addition, it is appreciated that diagnostic nucleotide sets that are developed for one use, e.g. to diagnose a particular disease, may later be found to be useful for a different application, e.g. to predict the likelihood that the particular disease will occur. Generally, the diagnostic nucleotide set will need to be validated for use in the second circumstance. As discussed herein, the sequence of diagnostic nucleotide set members may be amplified from RNA or cDNA using methods known in the art providing specific amplification of the nucleotide sequences.

Identification of novel nucleotide sequences that are differentially expressed in leukocytes

Novel nucleotide sequences that are differentially expressed in leukocytes are also part of the invention. Previously unidentified open reading frames may be identified in a library of differentially expressed candidate nucleotide sequences, as described above, and the DNA and predicted protein sequence may be identified and characterized as noted above. We identified unnamed (not previously described as corresponding to a gene, or an expressed gene) nucleotide sequences in the our candidate nucleotide library, depicted in Table 3A, 3B and the sequence listing. Accordingly, further embodiments of the invention are the isolated nucleic acids described in Tables 3A and 3B, and in the sequence listing. The novel differentially expressed nucleotide sequences of the invention are useful in the diagnostic nucleotide set of the invention described above, and are further useful as members of a diagnostic nucleotide set immobilized on an array. The novel partial nucleotide sequences may be further characterized using sequence tools and publically or privately accessible sequence databases, as is well known in the art: Novel differentially expressed nucleotide sequences may be identified as disease target nucleotide sequences, described below. Novel nucleotide sequences may also be used as imaging reagent, as further described below.

As used herein, "novel nucleotide sequence" refers to (a) a nucleotide sequence containing at least one of the DNA sequences disclosed herein (as shown in FIGS. Table 3A, 3B and the sequence listing); (b) any DNA sequence that encodes the amino acid sequence encoded by the DNA sequences disclosed herein; (c) any

DNA sequence that hybridizes to the complement of the coding sequences disclosed herein, contained within the coding region of the nucleotide sequence to which the DNA sequences disclosed herein (as shown in Table 3A, 3B and the sequence listing) belong, under highly stringent conditions, e.g., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65° C, and washing in 0.1XSSC/0.1% SDS at 68° C. (Ausubel F. M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3), (d) any DNA sequence that hybridizes to the complement of the coding sequences disclosed herein, (as shown in Table 3A, 3B and the sequence listing) contained within the coding region of the nucleotide sequence to which DNA sequences disclosed herein (as shown in TABLES 3A, 3B and the sequence listing) belong, under less stringent conditions, such as moderately stringent conditions, e.g., washing in 0.2XSSC/0.1% SDS at 42°C. (Ausubel et al., 1989, supra), yet which still encodes a functionally equivalent gene product; and/or (e) any DNA sequence that is at least 90% identical, at least 80% identical or at least 70% identical to the coding sequences disclosed herein (as shown in TABLES 3A, 3B and the sequence listing), wherein % identity is determined using standard algorithms known in the art.

The invention also includes nucleic acid molecules, preferably DNA molecules, that hybridize to, and are therefore the complements of, the DNA sequences (a) through (c), in the preceding paragraph. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("oligos"), highly stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C. (for 14-base oligos), 48°C. (for 17-base oligos), 55°C. (for 20-base oligos), and 60°C. (for 23-base oligos). These nucleic acid molecules may act as target nucleotide sequence antisense molecules, useful, for example, in target nucleotide sequence regulation and/or as antisense primers in amplification reactions of target nucleotide sequence nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for target nucleotide sequence regulation. Still further, such molecules may be used as components of diagnostic methods whereby the presence of a disease-causing allele, may be detected.

The invention also encompasses (a) DNA vectors that contain any of the foregoing coding sequences and/or their complements (i.e., antisense); (b) DNA expression vectors that contain any of the foregoing coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences; and (c) genetically engineered host cells that contain any of the foregoing coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory elements include but are not limited to inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. The invention includes fragments of any of the DNA sequences disclosed herein. Fragments of the DNA sequences may be at least 5, at least 10, at least 15, at least 19 nucleotides, at least 25 nucleotides, at least 50 nucleotides, at least 100 nucleotides, at least 200, at least 500, or larger.

In addition to the nucleotide sequences described above, homologues of such sequences, as may, for example be present in other species, may be identified and may be readily isolated, without undue experimentation, by molecular biological techniques well known in the art, as well as use of gene analysis tools described above, and e.g., in Example 4. Further, there may exist nucleotide sequences at other genetic loci within the genome that encode proteins which have extensive homology to one or more domains of such gene products. These nucleotide sequences may also be identified via similar techniques.

For example, the isolated differentially expressed nucleotide sequence may be labeled and used to screen a cDNA library constructed from mRNA obtained from the organism of interest. Hybridization conditions will be of a lower stringency when the cDNA library was derived from an organism different from the type of organism from which the labeled sequence was derived. Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions. Such low stringency conditions will be well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Springs Harbor Press, N.Y.; and Ausubel et al., 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Novel nucleotide products include those proteins encoded by the novel nucleotide sequences described, above. Specifically, novel gene products may include polypeptides encoded by the novel nucleotide sequences contained in the coding regions of the nucleotide sequences to which DNA sequences disclosed herein (in TABLES 3A, 3B and the sequence listing).

In addition, novel protein products of novel nucleotide sequences may include proteins that represent functionally equivalent gene products. Such an equivalent novel gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the novel nucleotide sequences described, above, but which result in a silent change, thus producing a functionally equivalent novel nucleotide sequence product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved.

For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent", as utilized herein, refers to a protein capable of exhibiting a substantially similar in vivo activity as the endogenous novel gene products encoded by the novel nucleotide described, above.

The novel gene products (protein products of the novel nucleotide sequences) may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the novel gene polypeptides and peptides of the invention by expressing nucleic acid encoding novel nucleotide sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing novel nucleotide sequence protein coding sequences and appropriate transcriptional/translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques and in vivo recombination/genetic recombination. See, for example, the techniques described in Sambrook et al., 1989, supra, and Ausubel et al., 1989, supra. Alternatively, RNA capable of encoding novel nucleotide sequence protein sequences may be chemically synthesized using, for example, synthesizers. See, for example,

the techniques described in "Oligonucleotide Synthesis", 1984, Gait, M. J. ed., IRL Press, Oxford.

A variety of host-expression vector systems may be utilized to express the novel nucleotide sequence coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the novel protein encoded by the novel nucleotide sequence of the invention in situ. These include but are not limited to microorganisms such as bacteria (e.g., E. coli, B. subtilis) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing novel nucleotide sequence protein coding sequences; yeast (e.g. Saccharomyces, Pichia) transformed with recombinant yeast expression vectors containing the novel nucleotide sequence protein coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the novel nucleotide sequence protein coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing novel nucleotide sequence protein coding sequences; or mammalian cell systems (e.g. COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5 K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the novel nucleotide sequence protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the E. coli expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the novel nucleotide sequence protein coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the likes of pGEX vectors may

also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target nucleotide sequence protein can be released from the GST moiety. Other systems useful in the invention include use of the FLAG epitope or the 6-HIS systems.

In an insect system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign nucleotide sequences. The virus grows in Spodoptera frugiperda cells. The novel nucleotide sequence coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of novel nucleotide sequence coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect Spodoptera frugiperda cells in which the inserted nucleotide sequence is expressed. (E.g., see Smith et al., 1983, J. Virol. 46: 584; Smith, U.S. Pat. No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the novel nucleotide sequence coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric nucleotide sequence may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing novel nucleotide sequence encoded protein in infected hosts. (E.g., See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted novel nucleotide sequence coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire novel nucleotide sequence, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the novel nucleotide sequence coding sequence is inserted, exogenous translational control signals, including,

perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., 1987, Methods in Enzymol. 153:516-544).

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the product of the nucleotide sequence in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the novel nucleotide sequence encoded protein may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express novel nucleotide sequence encoded protein. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the novel nucleotide sequence encoded protein.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk-, hgprt- or aprt- cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); and hygro, which confers resistance to hygromycin (Santerre, et al., 1984, Gene 30:147) genes.

An alternative fusion protein system allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88: 8972-8976). In this system, the nucleotide sequence of interest is subcloned into a vaccinia recombination plasmid such that the nucleotide sequence's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni.sup.2 +-nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

Where recombinant DNA technology is used to produce the protein encoded by the novel nucleotide sequence for such assay systems, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization and/or detection.

Indirect labeling involves the use of a protein, such as a labeled antibody, which specifically binds to the protein encoded by the novel nucleotide sequence. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by an Fab expression library.

The invention also provides for antibodies to the protein encoded by the novel nucleotide sequences. Described herein are methods for the production of antibodies capable of specifically recognizing one or more novel nucleotide sequence epitopes. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')2 fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

Such antibodies may be used, for example, in the detection of a novel nucleotide sequence in a biological sample, or, alternatively, as a method for the inhibition of abnormal gene activity, for example, the inhibition of a disease target nucleotide sequence, as further described below. Thus, such antibodies may be utilized as part of cardiovascular or other disease treatment method, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of novel nucleotide sequence encoded proteins, or for the presence of abnormal forms of the such proteins.

For the production of antibodies to a novel nucleotide sequence, various host animals may be immunized by injection with a novel protein encoded by the novel nucleotide sequence, or a portion thereof. Such host animals may include but are not limited to rabbits, mice, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as novel gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with novel gene product supplemented with adjuvants as also described above.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to the hybridoma technique of Kohler and Milstein, (1975, Nature 256:495-497; and U.S. Pat. No. 4,376,110), the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al., 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD

and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger et al., 1984, Nature, 312:604-608; Takeda et al., 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Pat. No. 4,946,778; Bird, 1988, Science 242:423-426; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-546) can be adapted to produce novel nucleotide sequence-single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques For example, such fragments include but are not limited to: the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Disease specific target nucleotide sequences

The invention also provides disease specific target nucleotide sequences, and sets of disease specific target nucleotide sequences. The diagnostic nucleotide sets, subsets thereof, novel nucleotide sequences, and individual members of the diagnostic nucleotide sets identified as described above are also disease specific target nucleotide sequences. In particular, individual nucleotide sequences that are differentially regulated or have predictive value that is strongly correlated with a disease or disease criterion are especially favorable as disease specific target nucleotide sequences. Sets of genes that are co-regulated may also be identified as disease specific target

nucleotide sets. Such nucleotide sequences and/or nucleotide sequence products are targets for modulation by a variety of agents and techniques. For example, disease specific target nucleotide sequences (or the products of such nucleotide sequences, or sets of disease specific target nucleotide sequences) can be inhibited or activated by, e.g., target specific monoclonal antibodies or small molecule inhibitors, or delivery of the nucleotide sequence or gene product of the nucleotide sequence to patients. Also, sets of genes can be inhibited or activated by a variety of agents and techniques. The specific usefulness of the target nucleotide sequence(s) depends on the subject groups from which they were discovered, and the disease or disease criterion with which they correlate.

Imaging

The invention also provides for imaging reagents. The differentially expressed leukocyte nucleotide sequences, diagnostic nucleotide sets, or portions thereof, and novel nucleotide sequences of the invention are nucleotide sequences expressed in cells with or without disease. Leukocytes expressing a nucleotide sequence(s) that is differentially expressed in a disease condition may localize within the body to sites that are of interest for imaging purposes. For example, a leukocyte expressing a nucleotide sequence(s) that are differentially expressed in an individual having atherosclerosis may localize or accumulate at the site of an atherosclerotic placque. Such leukocytes, when labeled, may provide a detection reagent for use in imaging regions of the body where labeled leukocyte accumulate or localize, for example, at the atherosclerotic plaque in the case of atherosclerosis. For example, leukocytes are collected from a subject, labeled in vitro, and reintroduced into a subject. Alternatively, the labeled reagent is introduced into the subject individual, and leukocyte labeling occurs within the patient.

Imaging agents that detect the imaging targets of the invention are produced by well-known molecular and immunological methods (for exemplary protocols, *see*, e.g., Ausubel, Berger, and Sambrook, as well as Harlow and Lane, *supra*).

For example, a full-length nucleic acid sequence, or alternatively, a gene fragment encoding an immunogenic peptide or polypeptide fragments, is cloned into a convenient expression vector, for example, a vector including an in-frame epitope or substrate binding tag to facilitate subsequent purification. Protein is then expressed from the cloned cDNA sequence and used to generate antibodies, or other specific

binding molecules, to one or more antigens of the imaging target protein. Alternatively, a natural or synthetic polypeptide (or peptide) or small molecule that specifically binds (or is specifically bound to) the expressed imaging target can be identified through well established techniques (*see*, e.g., Mendel et al. (2000) Anticancer Drug Des 15:29-41; Wilson (2000) Curr Med Chem 7:73-98; Hamby and Showwalter (1999) Pharmacol Ther 82:169-93; and Shimazawa et al. (1998) Curr Opin Struct Biol 8:451-8). The binding molecule, e.g., antibody, small molecule ligand, etc., is labeled with a contrast agent or other detectable label, e.g., gadolinium, iodine, or a gamma-emitting source. For in-vivo imaging of a disease process that involved leukocytes, the labeled antibody is infused into a subject, e.g., a human patient or animal subject, and a sufficient period of time is passed to permit binding of the antibody to target cells. The subject is then imaged with appropriate technology such as MRI (when the label is gadolinium) or with a gamma counter (when the label is a gamma emitter).

Identification of nucleotide sequence involved in leukocyte adhesion

The invention also encompasses a method of identifying nucleotide sequences involved in leukocyte adhesion. The interaction between the endothelial cell and leukocyte is a fundamental mechanism of all inflammatory disorders, including the diseases listed in Table 1. For example, the first visible abnormality in atherosclerosis is the adhesion to the endothelium and diapedesis of mononuclear cells (e.g., T-cell and monocyte). Insults to the endothelium (for example, cytokines, tobacco, diabetes, hypertension and many more) lead to endothelial cell activation. The endothelium then expresses adhesion molecules, which have counter receptors on mononuclear cells. Once the leukocyte receptors have bound the endothelial adhesion molecules, they stick to the endothelium, roll a short distance, stop and transmigrate across the endothelium. A similar set of events occurs in both acute and chronic inflammation

Human endothelial cells, e.g. derived from human coronary arteries, human aorta, human pulmonary artery, human umbilical vein or microvascular endothelial cells, are cultured as a confluent monolayer, using standard methods. Some of the endothelial cells are then exposed to cytokines or another activating stimuli such as oxidized LDL, hyperglycemia, shear stress, or hypoxia (Moser et al. 1992). Some endothelial cells are not exposed to such stimuli and serve as controls. For example, the endothelial cell monolayer is incubated with culture medium containing 5 U/ml of human recombinant IL-1alpha or 10 ng/ml TNF (tumor necrosis factor), for a period of minutes to overnight. The culture medium composition is changed or the flask is sealed to induce hypoxia. In addition, tissue culture plate is rotated to induce sheer stress.

Human T-cells and/or monocytes are cultured in tissue culture flasks or plates, with LGM-3 media from Clonetics. Cells are incubated at 37 degree C, 5% CO2 and 95% humidity. These leukocytes are exposed to the activated or control endothelial layer by adding a suspension of leukocytes on to the endothelial cell monolayer. The endothelial cell monolayer is cultured on a tissue culture treated plate/ flask or on a microporous membrane. After a variable duration of exposures, the endothelial cells and leukocytes are harvested separately by treating all cells with trypsin and then sorting the endothelial cells from the leukocytes by magnetic affinity reagents to an endothelial cell specific marker such as PECAM-1 (Stem Cell Technologies). RNA is extracted from the isolated cells by standard techniques. Leukocyte RNA is labeled as described above, and hybridized to leukocyte candidate nucleotide library. Epithelial cell RNA is also labeled and hybridized to the leukocyte candidate nucleotide library. Alternatively, the epithelial cell RNA is hybridized to a epithelial cell candidate nucleotide library, prepared according to the methods described for leukocyte candidate libraries, above.

Hybridization to candidate nucleotide libraries will reveal nucleotide sequences that are up-regulated or down-regulated in leukocyte and/or epithelial cells undergoing adhesion. The differentially regulated nucleotide sequences are further characterized, e.g. by isolating and sequencing the full-length sequence, analysis of the DNA and predicted protein sequence, and functional characterization of the protein product of the nucleotide sequence, as described above. Further characterization may result in the identification of leukocyte adhesion specific target nucleotide sequences, which may be candidate targets for regulation of the

inflammatory process. Small molecule or antibody inhibitors can be developed to inhibit the target nucleotide sequence function. Such inhibitors are tested for their ability to inhibit leukocyte adhesion in the in vitro test described above.

Integrated systems

Integrated systems for the collection and analysis of expression profiles, and molecular signatures, as well as for the compilation, storage and access of the databases of the invention, typically include a digital computer with software including an instruction set for sequence searching and analysis, and, optionally, high-throughput liquid control software, image analysis software, data interpretation software, a robotic control armature for transferring solutions from a source to a destination (such as a detection device) operably linked to the digital computer, an input device (e.g., a computer keyboard) for entering subject data to the digital computer, or to control analysis operations or high throughput sample transfer by the robotic control armature. Optionally, the integrated system further comprises an image scanner for digitizing label signals from labeled assay components, e.g., labeled nucleic acid hybridized to a candidate library microarray. The image scanner can interface with image analysis software to provide a measurement of the presence or intensity of the hybridized label, i.e., indicative of an on/off expression pattern or an increase or decrease in expression.

Readily available computational hardware resources using standard operating systems are fully adequate, e.g., a PC (Intel x86 or Pentium chip- compatible DOS,™ OS2,™ WINDOWS,™ WINDOWS NT,™ WINDOWS95,™ WINDOWS98,™ LINUX, or even Macintosh, Sun or PCs will suffice) for use in the integrated systems of the invention. Current art in software technology is similarly adequate (i.e., there are a multitude of mature programming languages and source code suppliers) for design, e.g., of an upgradeable open-architecture object-oriented heuristic algorithm, or instruction set for expression analysis, as described herein. For example, software for aligning or otherwise manipulating ,molecular signatures can be constructed by one of skill using a standard programming language such as Visual basic, Fortran, Basic, Java, or the like, according to the methods herein.

Various methods and algorithms, including genetic algorithms and neural networks, can be used to perform the data collection, correlation, and storage functions, as well as other desirable functions, as described herein. In addition, digital

or analog systems such as digital or analog computer systems can control a variety of other functions such as the display and/or control of input and output files.

For example, standard desktop applications such as word processing software (e.g., Corel WordPerfect™ or Microsoft Word™) and database software (e.g., spreadsheet software such as Corel Quattro Pro™, Microsoft Excel™, or database programs such as Microsoft Access™ or Paradox™) can be adapted to the present invention by inputting one or more character string corresponding, e.g., to an expression pattern or profile, subject medical or historical data, molecular signature, or the like, into the software which is loaded into the memory of a digital system, and carrying out the operations indicated in an instruction set, e.g., as exemplified in Figure 2. For example, systems can include the foregoing software having the appropriate character string information, e.g., used in conjunction with a user interface in conjunction with a standard operating system such as a Windows, Macintosh or LINUX system. For example, an instruction set for manipulating strings of characters, either by programming the required operations into the applications or with the required operations performed manually by a user (or both). For example, specialized sequence alignment programs such as PILEUP or BLAST can also be incorporated into the systems of the invention, e.g., for alignment of nucleic acids or proteins (or corresponding character strings).

Software for performing the statistical methods required for the invention, e.g., to determine correlations between expression profiles and subsets of members of the diagnostic nucleotide libraries, such as programmed embodiments of the statistical methods described above, are also included in the computer systems of the invention. Alternatively, programming elements for performing such methods as principle component analysis (PCA) or least squares analysis can also be included in the digital system to identify relationships between data. Exemplary software for such methods is provided by Partek, Inc., St. Peter, Mo; http://www.partek.com.

Any controller or computer optionally includes a monitor which can include, e.g., a flat panel display (e.g., active matrix liquid crystal display, liquid crystal display), a cathode ray tube ("CRT") display, or another display system which serves as a user interface, e.g., to output predictive data. Computer circuitry, including numerous integrated circuit chips, such as a microprocessor, memory, interface circuits, and the like, is often placed in a casing or box which optionally also includes

a hard disk drive, a floppy disk drive, a high capacity removable drive such as a writeable CD-ROM, and other common peripheral elements.

Inputting devices such as a keyboard, mouse, or touch sensitive screen, optionally provide for input from a user and for user selection, e.g., of sequences or data sets to be compared or otherwise manipulated in the relevant computer system. The computer typically includes appropriate software for receiving user instructions, either in the form of user input into a set parameter or data fields (e.g., to input relevant subject data), or in the form of preprogrammed instructions, e.g., preprogrammed for a variety of different specific operations. The software then converts these instructions to appropriate language for instructing the system to carry out any desired operation.

The integrated system may also be embodied within the circuitry of an application specific integrated circuit (ASIC) or programmable logic device (PLD). In such a case, the invention is embodied in a computer readable descriptor language that can be used to create an ASIC or PLD. The integrated system can also be embodied within the circuitry or logic processors of a variety of other digital apparatus, such as PDAs, laptop computer systems, displays, image editing equipment, etc.

The digital system can comprise a learning component where expression profiles, and relevant subject data are compiled and monitored in conjunction with physical assays, and where correlations, e.g., molecular signatures with predictive value for a disease, are established or refined. Successful and unsuccessful combinations are optionally documented in a database to provide justification/preferences for user-base or digital system based selection of diagnostic nucleotide sets with high predictive accuracy for a specified disease or condition.

The integrated systems can also include an automated workstation. For example, such a workstation can prepare and analyze leukocyte RNA samples by performing a sequence of events including: preparing RNA from a human blood sample; labeling the RNA with an isotopic or non-isotopic label; hybridizing the labeled RNA to at least one array comprising all or part of the candidate library; and detecting the hybridization pattern. The hybridization pattern is digitized and recorded in the appropriate database.

Automated RNA preparation tool

The invention also includes an automated RNA preparation tool for the preparation of mononuclear cells from whole blood samples, and preparation of RNA from the mononuclear cells. In a preferred embodiment, the use of the RNA preparation tool is fully automated, so that the cell separation and RNA isolation would require no human manipulations. Full automation is advantageous because it minimizes delay, and standardizes sample preparation across different laboratories. This standardization increases the reproducibility of the results.

Figure 2 depicts the processes performed by the RNA preparation tool of the invention. A primary component of the device is a centrifuge (A). Tubes of whole blood containing a density gradient solution, transcription/translation inhibitors, and a gel barrier that separates erythrocytes from mononuclear cells and serum after centrifugation are placed in the centrifuge (B). The barrier is permeable to erythrocytes and granulocytes during centrifugation, but does not allow mononuclear cells to pass through (or the barrier substance has a density such that mononuclear cells remain above the level of the barrier during the centrifugation). After centrifugation, the erythrocytes and granulocytes are trapped beneath the barrier, facilitating isolation of the mononuclear cell and serum layers. A mechanical arm removes the tube and inverts it to mix the mononuclear cell layer and the serum (C). The arm next pours the supernatant into a fresh tube (D), while the erythrocytes and granulocytes remained below the barrier. Alternatively, a needle is used to aspirate the supernatant and transfer it to a fresh tube. The mechanical arms of the device opens and closes lids, dispenses PBS to aid in the collection of the mononuclear cells by centrifugation, and moves the tubes in and out of the centrifuge. Following centrifugation, the supernatant is poured off or removed by a vacuum device (E), leaving an isolated mononuclear cell pellet. Purification of the RNA from the cells is performed automatically, with lysis buffer and other purification solutions (F) automatically dispensed and removed before and after centrifugation steps. The result is a purified RNA solution. In another embodiment, RNA isolation is performed using a column or filter method. In yet another embodiment, the invention includes an on-board homogenizer for use in cell lysis.

Other automated systems

Automated and/or semi-automated methods for solid and liquid phase high-throughput sample preparation and evaluation are available, and supported by commercially available devices. For example, robotic devices for preparation of nucleic acids from bacterial colonies, e.g., to facilitate production and characterization of the candidate library include, for example, an automated colony picker (e.g., the Q-bot, Genetix, U.K.) capable of identifying, sampling, and inoculating up to 10,000/4 hrs different clones into 96 well microtiter dishes. Alternatively, or in addition, robotic systems for liquid handling are available from a variety of sources, e.g., automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Beckman Coulter, Inc. (Fullerton, CA)) which mimic the manual operations performed by a scientist. Any of the above devices are suitable for use with the present invention, e.g., for high-throughput analysis of library components or subject leukocyte samples. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art.

High throughput screening systems that automate entire procedures, e.g., sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the relevant assay are commercially available. (*see, e.g.,* Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, *etc.*). These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. Similarly, arrays and array readers are available, e.g., from Affymetrix, PE Biosystems, and others.

The manufacturers of such systems provide detailed protocols the various high throughput. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

A variety of commercially available peripheral equipment, including, e.g., optical and fluorescent detectors, optical and fluorescent microscopes, plate readers, CCD arrays, phosphorimagers, scintillation counters, phototubes, photodiodes, and the like, and software is available for digitizing, storing and analyzing a digitized video or digitized optical or other assay results, e.g., using PC (Intel x86 or pentium

chip- compatible DOS™, OS2™ WINDOWS™, WINDOWS NT™ or WINDOWS95™ based machines), MACINTOSH™, or UNIX based (*e.g.*, SUN™ work station) computers.

Embodiment in a web site.

The methods described above can be implemented in a localized or distributed computing environment. For example, if a localized computing environment is used, an array comprising a candidate nucleotide library, or diagnostic nucleotide set, is configured in proximity to a detector, which is, in turn, linked to a computational device equipped with user input and output features.

In a distributed environment, the methods can be implemented on a single computer with multiple processors or, alternatively, on multiple computers. The computers can be linked, e.g. through a shared bus, but more commonly, the computer(s) are nodes on a network. The network can be generalized or dedicated, at a local level or distributed over a wide geographic area. In certain embodiments, the computers are components of an intra-net or an internet.

The predictive data corresponding to subject molecular signatures (e.g., expression profiles, and related diagnostic, prognostic, or monitoring results) can be shared by a variety of parties. In particular, such information can be utilized by the subject, the subject's health care practitioner or provider, a company or other institution, or a scientist. An individual subject's data, a subset of the database or the entire database recorded in a computer readable medium can be accessed directly by a user by any method of communication, including, but not limited to, the internet. With appropriate computational devices, integrated systems, communications networks, users at remote locations, as well as users located in proximity to, e.g., at the same physical facility, the database can access the recorded information. Optionally, access to the database can be controlled using unique alphanumeric passwords that provide access to a subset of the data. Such provisions can be used, e.g., to ensure privacy, anonymity, etc.

Typically, a client (e.g., a patient, practitioner, provider, scientist, or the like) executes a Web browser and is linked to a server computer executing a Web server. The Web browser is, for example, a program such as IBM's Web Explorer, Internet explorer, NetScape or Mosaic, or the like. The Web server is typically, but not necessarily, a program such as IBM's HTTP Daemon or other WWW daemon (e.g.,

LINUX-based forms of the program). The client computer is bi-directionally coupled with the server computer over a line or via a wireless system. In turn, the server computer is bi-directionally coupled with a website (server hosting the website) providing access to software implementing the methods of this invention.

A user of a client connected to the Intranet or Internet may cause the client to request resources that are part of the web site(s) hosting the application(s) providing an implementation of the methods described herein. Server program(s) then process the request to return the specified resources (assuming they are currently available). A standard naming convention has been adopted, known as a Uniform Resource Locator ("URL"). This convention encompasses several types of location names, presently including subclasses such as Hypertext Transport Protocol ("http"), File Transport Protocol ("ftp"), gopher, and Wide Area Information Service ("WAIS"). When a resource is downloaded, it may include the URLs of additional resources. Thus, the user of the client can easily learn of the existence of new resources that he or she had not specifically requested.

Methods of implementing Intranet and/or Intranet embodiments of computational and/or data access processes are well known to those of skill in the art and are documented, e.g., in ACM Press, pp. 383-392; ISO-ANSI, Working Draft, "Information Technology-Database Language SQL", Jim Melton, Editor, International Organization for Standardization and American National Standards Institute, Jul. 1992; ISO Working Draft, "Database Language SQL-Part 2:Foundation (SQL/Foundation)", CD9075-2:199.chi.SQL, Sep. 11, 1997; and Cluer et al. (1992) A General Framework for the Optimization of Object-Oriented Queries, Proc SIGMOD International Conference on Management of Data, San Diego, California, Jun. 2-5, 1992, SIGMOD Record, vol. 21, Issue 2, Jun., 1992; Stonebraker, M., Editor;. Other resources are available, e.g., from Microsoft, IBM, Sun and other software development companies.

Using the tools described above, users of the reagents, methods and database as discovery or diagnostic tools can query a centrally located database with expression and subject data. Each submission of data adds to the sum of expression and subject information in the database. As data is added, a new correlation statistical analysis is automatically run that incorporates the added clinical and expression data. Accordingly, the predictive accuracy and the types of correlations of the recorded molecular signatures increases as the database grows.

For example, subjects, such as patients, can access the results of the expression analysis of their leukocyte samples and any accrued knowledge regarding the likelihood of the patient's belonging to any specified diagnostic (or prognostic, or monitoring, or risk group), i.e., their expression profiles, and/or molecular signatures. Optionally, subjects can add to the predictive accuracy of the database by providing additional information to the database regarding diagnoses, test results, clinical or other related events that have occurred since the time of the expression profiling. Such information can be provided to the database via any form of communication, including, but not limited to, the internet. Such data can be used to continually define (and redefine) diagnostic groups. For example, if 1000 patients submit data regarding the occurrence of myocardial infarction over the 5 years since their expression profiling, and 300 of these patients report that they have experienced a myocardial infarction and 700 report that they have not, then the 300 patients define a new "group A." As the algorithm is used to continually query and revise the database, a new diagnostic nucleotide set that differentiates groups A and B (i.e., with and without myocardial infarction within a five year period) is identified. This newly defined nucleotide set is then be used (in the manner described above) as a test that predicts the occurrence of myocardial infarction over a five-year period. While submission directly by the patient is exemplified above, any individual with access and authority to submit the relevant data e.g., the patient's physician, a laboratory technician, a health care or study administrator, or the like, can do so.

As will be apparent from the above examples, transmission of information via the internet (or via an intranet) is optionally bi-directional. That is, for example, data regarding expression profiles, subject data, and the like are transmitted via a communication system to the database, while information regarding molecular signatures, predictive analysis, and the like, are transmitted from the database to the user. For example, using appropriate configurations of an integrated system including a microarray comprising a diagnostic nucleotide set, a detector linked to a computational device can directly transmit (locally or from a remote workstation at great distance, e.g., hundreds or thousands of miles distant from the database) expression profiles and a corresponding individual identifier to a central database for analysis according to the methods of the invention. According to, e.g., the algorithms described above, the individual identifier is assigned to one or more diagnostic (or prognostic, or monitoring, etc.) categories. The results of this classification are then

relayed back, via, e.g., the same mode of communication, to a recipient at the same or different internet (or intranet) address.

Kits

The present invention is optionally provided to a user as a kit. Typically, a kit contains one or more diagnostic nucleotide sets of the invention. Alternatively, the kit contains the candidate nucleotide library of the invention. Most often, the kit contains a diagnostic nucleotide probe set, or other subset of a candidate library, e.g., as a cDNA or antibody microarray packaged in a suitable container. The kit may further comprise, one or more additional reagents, e.g., substrates, labels, primers, for labeling expression products, tubes and/or other accessories, reagents for collecting blood samples, buffers, e.g., erythrocyte lysis buffer, leukocyte lysis buffer, hybridization chambers, cover slips, etc., as well as a software package, e.g., including the statistical methods of the invention, e.g., as described above, and a password and/or account number for accessing the compiled database. The kit optionally further comprises an instruction set or user manual detailing preferred methods of using the diagnostic nucleotide sets in the methods of the invention. Exemplary kits are described in Figure 3.

This invention will be better understood by reference to the following non-limiting Examples:

EXAMPLES

List of Example titles

Example 1: Generation of subtracted leukocyte candidate nucleotide library

Example 2: Identification of nucleotide sequences for candidate library using data mining techniques

Example 3: DNA Sequencing and Processing of raw sequence data.

Example 4: Further sequence analysis of novel nucleotide sequences identified by subtractive hybridization screening

Example 5: Further sequence analysis of novel Clone 596H6

Example 6: Further sequence analysis of novel Clone 486E11

Example 7: Preparation of a leukocyte cDNA array comprising a candidate gene library

Example 8: Preparation of RNA from mononuclear cells for expression profiling

Example 9: Preparation of Buffy Coat Control RNA for use in leukocyte expression profiling

Example 10. RNA Labeling and hybridization to a leukocyte cDNA array of candidate nucleotide sequences.

Example 11: Identification of diagnostic gene sets useful in diagnosis and treatment of Cardiac allograft rejection

Example 12: Identification of diagnostic nucleotide sets for kidney and liver allograft rejection

Example 13: Identification of diagnostic nucleotide sequences sets for use in the diagnosis and treatment of Atherosclerosis, Stable Angina Pectoris, and acute coronary syndrome.

Example 14: Identification of diagnostic nucleotide sets for use in diagnosing and treating Restenosis

Example 15: Identification of diagnostic nucleotide sets for use in monitoring treatment and/or progression of Congestive Heart Failure

Example 16: Identification of diagnostic nucleotide sets for use in diagnosis of rheumatoid arthritis.

Example 17: Identification of diagnostic nucleotide sets for diagnosis of cytomegalovirus

Example 18: Identification of diagnostic nucleotide sets for diagnosis of Epstein Barr Virus

Example 19: Identification of diagnostic nucleotides sets for monitoring response to statin drugs.

Example 20: Probe selection for a 24,000 feature Array.

Example 21: Design of oligonucleotide probes.

Example 22: Production of an array of 8,000 spotted 50 mer oligonucleotides.

Example 23: Amplification, labeling and hybridization of total RNA to an oligonucleotide microarray.

Example 24:Analysis of Human Transplant Patient Mononuclear cell RNA Hybridized to a 24,000 Feature Microarray.

Examples

Example 1: Generation of subtracted leukocyte candidate nucleotide library

To produce a candidate nucleotide library with representatives from the spectrum of nucleotide sequences that are differentially expressed in leukocytes, subtracted hybridization libraries were produced from the following cell types and conditions:

1. Buffy Coat leukocyte fractions - stimulated with ionomycin and PMA
2. Buffy Coat leukocyte fractions – un-stimulated
3. Peripheral blood mononuclear cells – stimulated with ionomycin and PMA
4. Peripheral blood mononuclear cells – un-stimulated
5. T lymphocytes – stimulated with PMA and ionomycin
6. T lymphocytes – resting

Cells were obtained from multiple individuals to avoid introduction of bias by using only one person as a cell source.

Buffy coats (platelets and leukocytes that are isolated from whole blood) were purchased from Stanford Medical School Blood Center. Four buffy coats were used, each of which was derived from about 350 ml of whole blood from one donor individual 10 ml of buffy coat sample was drawn from the sample bag using a needle and syringe. 40 ml of Buffer EL (Qiagen) was added per 10 ml of buffy coat to lyse red blood cells. The sample was placed on ice for 15 minutes, and cells were collected by centrifugation at 2000 rpm for 10 minutes. The supernatant was decanted and the cell pellet was re-suspended in leukocyte growth media supplemented with DNase (LGM-3 from Clonetics supplemented with Dnase at a final concentration of 30 U/ml). Cell density was determined using a hemocytometer. Cells were plated in media at a density of $1x10^6$ cells/ml in a total volume of 30 ml in a T-75 flask (Corning). Half of the cells were stimulated with ionomycin and phorbol myristate acetate (PMA) at a final concentration of 1 μg/ml and 62 ng/ml, respectively. Cells were incubated at 37°C and at 5% CO_2 for 3 hours, then cells were scraped off the flask and collected into 50 ml tubes. Stimulated and resting cell populations were kept separate. Cells were centrifuged at 2000 rpm for 10 minutes and the supernatant was removed. Cells were lysed in 6 ml of phenol/guanidine isothyocyanate (Trizol reagent, GibcoBRL), homogenized using a rotary

homogenizer, and frozen at 80°. Total RNA and mRNA were isolated as described below.

Two frozen vials of $5x10^6$ human peripheral blood mononuclear cells (PBMCs) were purchased from Clonetics (catalog number cc-2702). The cells were rapidly thawed in a 37°C water bath and transferred to a 15 ml tube containing 10 ml of leukocyte growth media supplemented with DNase (prepared as described above). Cells were centrifuged at 200μg for 10 minutes. The supernatant was removed and the cell pellet was resuspended in LGM-3 media supplemented with DNase. Cell density was determined using a hemocytometer. Cells were plated at a density of $1x10^6$ cells/ml in a total volume of 30 ml in a T-75 flask (Corning). Half of the cells were stimulated with ionomycin and PMA at a final concentration of 1 μg/ml and 62 ng/ml, respectively. Cells were incubated at 37°C and at 5% CO_2 for 3 hours, then cells were scraped off the flask and collected into 50 ml tubes. Stimulated and resting cell populations were kept separate. Cells were centrifuged at 2000 rpm and the supernatant was removed. Cells were lysed in 6 ml of phenol/guanidine isothyocyanate solution (TRIZOL reagent, GibcoBRL)), homogenized using a rotary homogenizer, and frozen at 80°. Total RNA and mRNA were isolated from these samples using the protocol described below.

45 ml of whole blood was drawn from a peripheral vein of four healthy human subjects into tubes containing anticoagulant. 50 μl RosetteSep (Stem Cell Technologies) cocktail per ml of blood was added, mixed well, and incubated for 20 minutes at room temperature. The mixture was diluted with an equal volume of PBS + 2% fetal bovine serum (FBS) and mixed by inversion. 30 ml of diluted mixture sample was layered on top of 15 ml DML medium (Stem Cell Technologies). The sample tube was centrifuged for 20 minutes at 1200xg at room temperature. The enriched T-lymphocyte cell layer at the plasma : medium interface was removed. Enriched cells were washed with PBS + 2% FBS and centrifuged at 1200 x g. The cell pellet was treated with 5 ml of erythrocyte lysis buffer (EL buffer, Qiagen) for 10 minutes on ice. The sample was centrifuged for 5 min at 1200g. Cells were plated at a density of $1x10^6$ cells/ml in a total volume of 30 ml in a T-75 flask (Corning). Half of the cells were stimulated with ionomycin and PMA at a final concentration of 1 μg/ml and 62 ng/ml, respectively. Cells were incubated at 37°C and at 5% CO_2 for 3 hours, then cells were scraped off the flask and collected into 50 ml tubes. Stimulated and resting cell populations were kept separate. Cells were centrifuged at 2000 rpm

and the supernatant was removed. Cells were lysed in 6 ml of phenol/guanidine isothyocyanate solution (TRIZOL reagent, GibcoBRL), homogenized using a rotary homogenizer, and frozen at 80°. Total RNA and mRNA were isolated as described below.

Total RNA and mRNA were isolated using the following procedure: the homogenized samples were thawed and mixed by vortexing. Samples were lysed in a 1:0.2 mixture of Trizol and chloroform, respectively. For some samples, 6 ml of Trizol-chloroform was added. Variable amounts of Trizol-chloroform was added to other samples. Following lysis, samples were centrifuged at 3000 g for 15 min at 4°C. The aqueous layer was removed into a clean tube and 4 volumes of Buffer RLT Qiagen) was added for every volume of aqueous layer. The samples were mixed thoroughly and total RNA was prepared from the sample by following the Qiagen Rneasy midi protocol for RNA cleanup (October 1999 protocol, Qiagen). For the final step, the RNA was eluted from the column twice with 250 µl Rnase-free water. Total RNA was quantified using a spectrophotometer. Isolation of mRNA from total RNA sample was done using The Oligotex mRNA isolation protocol (Qiagen) was used to isolate mRNA from total RNA, according to the manufacturer's instructions (Qiagen, 7/99 version). mRNA was quantified by spectrophotometry.

Subtracted cDNA libraries were prepared using Clontech's PCR-Select cDNA Subtraction Kit (protocol number PT-1117-1) as described in the manufacturer's protocol. The protocol calls for two sources of RNA per library, designated "Driver" and "Tester." The following 6 libraries were made:

Library	Driver RNA	Tester RNA
Buffy Coat Stimulated	Un-stimulated Buffy Coat	Stimulated Buffy Coat
Buffy Coat Resting	Stimulated Buffy Coat	Un-stimulated Buffy Coat
PBMC Stimulated	Un-stimulated PBMCs	Stimulated PBMCs
PBMC Resting	Stimulated PBMCs	Un-stimulated PBMCs
T-cell Stimulated	Un-stimulated T-cells	Stimulated T-cells
T-cell Resting	Stimulated T-cells	Un-stimulated T-cells

The Clontech protocol results in the PCR amplification of cDNA products. The PCR products of the subtraction protocol were ligated to the pGEM T-easy bacterial vector as described by the vector manufacturer (Promega 6/99 version). Ligated vector was transformed into competent bacteria using well-known techniques,

plated, and individual clones are picked, grown and stored as a glycerol stock at – 80C. Plasmid DNA was isolated from these bacteria by standard techniques and used for sequence analysis of the insert. Unique cDNA sequences were searched in the Unigene database (build 133), and Unigene cluster numbers were identified that corresponded to the DNA sequence of the cDNA. Unigene cluster numbers were recorded in an Excel spreadsheet.

Example 2: Identification of nucleotide sequences for candidate library using data mining techniques

Existing and publicly available gene sequence databases were used to identify candidate nucleotide sequences for leukocyte expression profiling. Genes and nucleotide sequences with specific expression in leukocytes, for example, lineage specific markers, or known differential expression in resting or activated leukocytes were identified. Such nucleotide sequences are used in a leukocyte candidate nucleotide library, alone or in combination with nucleotide sequences isolated through cDNA library construction, as described above.

Leukocyte candidate nucleotide sequences were identified using three primary methods. First, the publically accessible publication database PubMed was searched to identify nucleotide sequences with known specific or differential expression in leukocytes. Nucleotide sequences were identified that have been demonstrated to have differential expression in peripheral blood leukocytes between subjects with and without particular disease(s) selected from Table 1. Additionally, genes and gene sequences that were known to be specific or selective for leukocytes or sub-populations of leukocytes were identified in this way.

Next, two publicly available databases of DNA sequences, Unigene (http://www.ncbi.nlm.nih.gov/UniGene/) and BodyMap (http://bodymap.ims.u-tokyo.ac.jp/), were searched for sequenced DNA clones that showed specificity to leukocyte lineages, or subsets of leukocytes, or resting or activated leukocytes.

The human Unigene database (build 133) was used to identify leukocyte candidate nucleotide sequences that were likely to be highly or exclusively expressed in leukocytes. We used the Library Differential Display utility of Unigene (http://www.ncbi.nlm.nih.gov/UniGene/info/ddd.html), which uses statistical methods (The Fisher Exact Test) to identify nucleotide sequences that have relative specificity

for a chosen library or group of libraries relative to each other. We compared the following human libraries from Unigene release 133:

546 NCI_CGAP_HSC1 (399)
848 Human_mRNA_from_cd34+_stem_cells (122)
105 CD34+DIRECTIONAL (150)
3587 KRIBB_Human_CD4_intrathymic_T-cell_cDNA_library (134)
3586 KRIBB_Human_DP_intrathymic_T-cell_cDNA_library (179)
3585 KRIBB_Human_TN_intrathymic_T-cell_cDNA_library (127)
3586 323 Activated_T-cells_I (740)
376 Activated_T-cells_XX (1727)
327 Monocytes,_stimulated_II (110)
824 Proliferating_Erythroid_Cells_(LCB:ad_library) (665)
825 429 Macrophage_II (105)
387 Macrophage_I (137)
669 NCI_CGAP_CLL1 (11626)
129 Human_White_blood_cells (922)
1400 NIH_MGC_2 (422)
55 Human_promyelocyte (1220)
1010 NCI_CGAP_CML1 (2541)
2217 NCI_CGAP_Sub7 (218)
1395 NCI_CGAP_Sub6 (2764)
4874 NIH_MGC_48 (2524)

BodyMap, like Unigene, contains cell-specific libraries that contain potentially useful information about genes that may serve as lineage-specific or leukocyte specific markers (Okubo et al. 1992). We compared three leukocyte specific libraries, Granulocyte, CD4 T cell, and CD8 T cell , with the other libraries. Nucleotide sequences that were found in one or more of the leukocyte-specific libraries, but absent in the others, were identified. Clones that were found exclusively in one of the three leukocyte libraries were also included in a list of nucleotide sequences that could serve as lineage-specific markers.

Next, the sequence of the nucleotide sequences identified in PubMed or BodyMap were searched in Unigene (version 133), and a human Unigene cluster number was identified for each nucleotide sequence. The cluster number was

recorded in a Microsoft Excel™ spreadsheet, and a non-redundant list of these clones was made by sorting the clones by UniGene number, and removing all redundant clones using Microsoft Excel™ tools. The non-redundant list of UniGene cluster numbers was then compared to the UniGene cluster numbers of the cDNAs identified using differential cDNA hybridization, as described above in Example 1 (listed in Table 3 and the sequence listing). Only UniGene clusters that were not contained in the cDNA libraries were retained. Unigene clusters corresponding to 1911 candidate nucleotide sequences for leukocyte expression profiling were identified in this way and are listed in Table 3 and the sequence listing.

DNA clones corresponding to each UniGene cluster number are obtained in a variety of ways. First, a cDNA clone with identical sequence to part of, or all of the identified UniGene cluster is bought from a commercial vendor or obtained from the IMAGE consortium (http://image.llnl.gov/, the Integrated Molecular Analysis of Genomes and their Expression). Alternatively, PCR primers are designed to amplify and clone any portion of the nucleotide sequence from cDNA or genomic DNA using well-known techniques. Alternatively, the sequences of the identified UniGene clusters are used to design and synthesize oligonucleotide probes for use in microarray based expression profiling.

Example 3: DNA Sequencing and Processing of raw sequence data.

Clones of differentially expressed cDNAs (identified by subtractive hybridization, described above) were sequenced on an MJ Research BaseStation™ slab gel based fluorescent detection system, using BigDye™ (Applied Biosystems, Foster City, CA) terminator chemistry was used (Heiner et al., Genome Res 1998 May;8(5):557-61).

The fluorescent profiles were analyzed using the Phred sequence analysis program (Ewing et al, (1998), Genome Research 8: 175-185). Analysis of each clone results in a one pass nucleotide sequence and a quality file containing a number for each base pair with a score based on the probability that the determined base is correct. Each sequence files and its respective quality files were initially combined into single fasta format (Pearson, WR. Methods Mol Biol. 2000;132:185-219), multi-sequence file with the appropriate labels for each clone in the headers for subsequent automated analysis.

Initially, known sequences were analyzed by pair wise similarity searching using the blastn option of the blastall program obtained from the National Center for Biological Information, National Library of Medicine, National Institutes of Health (NCBI) to determine the quality score that produced accurate matching (Altschul SF,et al. J Mol Biol. 1990 Oct 5;215(3):403-10.). Empirically, it was determined that a raw score of 8 was the minimum that contained useful information. Using a sliding window average for 16 base pairs, an average score was determined. The sequence was removed (trimmed) when the average score fell below 8. Maximum reads were 950 nucleotides long.

Next, the sequences were compared by similarity matching against a database file containing the flanking vector sequences used to clone the cDNA, using the blastall program with the blastn option. All regions of vector similarity were removed, or "trimmed" from the sequences of the clones using scripts in the GAWK programming language, a variation of AWK (Aho AV et al, The Awk Programming Language (Addison-Wesley, Reading MA, 1988); Robbins, AD, "Effective AWK Programming" (Free Software Foundation, Boston MA, 1997). It was found that the first 45 base pairs of all the sequences were related to vector; these sequences were also trimmed and thus removed from consideration. The remaining sequences were then compared against the NCBI vector database (Kitts, P.A. et al. National Center for Biological Information, National Library of Medicine, National Institutes of Health, Manuscript in preparation (2001) using blastall with the blastn option. Any vector sequences that were found were removed from the sequences.

Messenger RNA contains repetitive elements that are found in genomic DNA. These repetitive elements lead to false positive results in similarity searches of query mRNA sequences versus known mRNA and EST databases. Additionally, regions of low information content (long runs of the same nucleotide, for example) also result in false positive results. These regions were masked using the program RepeatMasker2 found at http://repeatmasker.genome.washington.edu (Smit, AFA & Green, P "RepeatMasker" at http://ftp.genome.washington.edu/RM/RepeatMasker.html). The trimmed and masked files were then subjected to further sequence analysis.

Example 4: Further sequence analysis of novel nucleotide sequences identified by subtractive hybridization screening

cDNA sequences were further characterized using BLAST analysis. The BLASTN program was used to compare the sequence of the fragment to the UniGene, dbEST, and nr databases at NCBI (GenBank release 123.0; see Table 5). In the BLAST algorithm, the expect value for an alignment is used as the measure of its significance. First, the cDNA sequences were compared to sequences in Unigene (http://www.ncbi.nlm.nih.gov/UniGene). If no alignments were found with an expect value less than 10^{-25}, the sequence was compared to the sequences in the dbEST database using BLASTN. If no alignments were found with an expect value less than 10^{-25}, the sequence was compared to sequences in the nr database.

The BLAST analysis produced the following categories of results: a) a significant match to a known or predicted human gene, b) a significant match to a nonhuman DNA sequence, such as vector DNA or *E. coli* DNA, c) a significant match to an unidentified GenBank entry (a sequence not previously identified or predicted to be an expressed sequence or a gene), such as a cDNA clone, mRNA, or cosmid , or d) no significant alignments. If a match to a known or predicted human gene was found, analysis of the known or predicted protein product was performed as described below. If a match to an unidentified GenBank entry was found, or if no significant alignments were found, the sequence was searched against all known sequences in the human genome database (http://www.ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsBlast.html&&ORG=Hs, see Table 5).

If many unknown sequences were to be analyzed with BLASTN, the clustering algorithm CAP2 (Contig Assembly Program, version 2) was used to cluster them into longer, contiguous sequences before performing a BLAST search of the human genome. Sequences that can be grouped into contigs are likely to be cDNA from expressed genes rather than vector DNA, *E. coli* DNA or human chromosomal DNA from a noncoding region, any of which could have been incorporated into the library. Clustered sequences provide a longer query sequence for database comparisons with BLASTN, increasing the probability of finding a significant match to a known gene. When a significant alignment was found, further analysis of the putative gene was performed, as described below. Otherwise, the sequence of the

original cDNA fragment or the CAP2 contig is used to design a probe for expression analysis and further approaches are taken to identify the gene or predicted gene that corresponds to the cDNA sequence, including similarity searches of other databases, molecular cloning, and Rapid Amplification of cDNA Ends (RACE).

In some cases, the process of analyzing many unknown sequences with BLASTN was automated by using the BLAST network-client program blastcl3, which was downloaded from ftp://ncbi.nlm.nih.gov/blast/network/netblast.

When a cDNA sequence aligned to the sequence of one or more chromosomes, a large piece of the genomic region around the loci was used to predict the gene containing the cDNA. To do this, the contig corresponding to the mapped locus, as assembled by the RefSeq project at NCBI, was downloaded and cropped to include the region of alignment plus 100,000 bases preceding it and 100,000 bases following it on the chromosome. The result was a segment 200 kb in length, plus the length of the alignment. This segment, designated a putative gene, was analyzed using an exon prediction algorithm to determine whether the alignment area of the unknown sequence was contained within a region predicted to be transcribed (see Table 6).

This putative gene was characterized as follows: all of the exons comprising the putative gene and the introns between them were taken as a unit by noting the residue numbers on the 200kb+ segment that correspond to the first base of the first exon and the last base of the last exon, as given in the data returned by the exon prediction algorithm. The truncated sequence was compared to the UniGene, dbEST, and nr databases to search for alignments missed by searching with the initial fragment.

The predicted amino acid sequence of the gene was also analyzed. The peptide sequence of the gene predicted from the exons was used in conjunction with numerous software tools for protein analysis (see Table 7). These were used to classify or identify the peptide based on similarities to known proteins, as well as to predict physical, chemical, and biological properties of the peptides, including secondary and tertiary structure, flexibility, hydrophobicity, antigenicity (hydrophilicity), common domains and motifs, and localization within the cell or tissues. The peptide sequence was compared to protein databases, including SWISS-PROT, TrEMBL, GenPept, PDB, PIR, PROSITE, ProDom, PROSITE, Blocks,

PRINTS, and Pfam, using BLASTP and other algorithms to determine similarities to known proteins or protein subunits.

Example 5: Further sequence analysis of novel Clone 596H6

The sequence of clone 596H6 is provided below:

ACTATATTTA	GGCACCACTG	CCATAAACTA	CCAAAAAAAA	AATGTAATTC	50
CTAGAAGCTG	TGAAGAATAG	TAGTGTAGCT	AAGCACGGTG	TGTGGACAGT	10(
GGGACATCTG	CCACCTGCAG	TAGGTCTCTG	CACTCCCAAA	AGCAAATTAC	15(
ATTGGCTTGA	ACTTCAGTAT	GCCCGGTTCC	ACCCTCCAGA	AACTTTTGTG	20(
TTCTTTGTAT	AGAATTTAGG	AACTTCTGAG	GGCCACAAAT	ACACACATTA	25(
AAAAAGGTAG	AATTTTTGAA	GATAAGATTC	TTCTAAAAAA	GCTTCCCAAT	30(
GCTTGAGTAG	AAAGTATCAG	TAGAGGTATC	AAGGGAGGAG	AGACTAGGTG	35(
ACCACTAAAC	TCCTTCAGAC	TCTTAAAATT	ACGATTCTTT	TCTCAAAGGG	40(
GAAGAACGTC	AGTGCAGCGA	TCCCTTCACC	TTTAGCTAAA	GAATTGGACT	45(
GTGCTGCTCA	AAATAAAGAT	CAGTTGGAGG	TANGATGTCC	AAGACTGAAG	50(
GTAAAGGACT	AGTGCAAACT	GAAAGTGATG	GGGAAACAGA	CCTACGTATG	55(
GAAGCCATGT	AGTGTTCTTC	ACAGGCTGCT	GTTGACTGAA	ATTCCTATCC	60(
TCAAATTACT	CTAGACTGAA	GCTGCTTCCC	TTCAGTGAGC	AGCCTCTCCT	65(
TCCAAGATTC	TGGAAAGCAC	ACCTGACTCC	AAACAAAGAC	TTAGAGCCCT	70(
GTGTCAGTGC	TGCTGCTGCT	TTTACCAGAT	TCTCTAACCT	TCCGGGTAGA	75(
AGAG	(SEQ ID NO: 8767)				

This sequence was used as input for a series of BLASTN searches. First, it was used to search the UniGene database, build 132 (http://www.ncbi.nlm.nih.gov/BLAST/). No alignments were found with an expect value less than the threshold value of 10^{-25}. A BLASTN search of the database dbEST, release 041001, was then performed on the sequence and 21 alignments were found (http://www.ncbi.nlm.nih.gov/BLAST/). Ten of these had expect values less than 10^{-25}, but all were matches to unidentified cDNA clones. Next, the sequence was used to run a BLASTN search of the nr database, release 123.0. No significant alignment to any sequence in nr was found. Finally, a BLASTN search of the human genome was performed on the sequence (http://www.ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsBlast.html&&ORG=Hs).

A single alignment to the genome was found on contig NT_004698.3 (e=0.0). The region of alignment on the contig was from base 1,821,298 to base 1,822,054,

and this region was found to be mapped to chromosome 1, from base 105,552,694 to base 105,553,450. The sequence containing the aligned region, plus 100 kilobases on each side of the aligned region, was downloaded. Specifically, the sequence of chromosome 1 from base105,452,694 to 105,653,450 was downloaded (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/seq_reg.cgi?chr=1&from=105452694&to=105653450).

This 200,757 bp segment of the chromosome was used to predict exons and their peptide products as follows. The sequence was used as input for the Genscan algorithm (http://genes.mit.edu/GENSCAN.html), using the following Genscan settings:

Organism: vertebrate

Suboptimal exon cutoff: 1.00 (no suboptimal exons)

Print options: Predicted CDS and peptides

The region matching the sequence of clone 596H6 was known to span base numbers 100,001 to 100,757 of the input sequence. An exon was predicted by the algorithm, with a probability of 0.695, covering bases 100,601 to 101,094 (designated exon 4.14 of the fourth predicted gene). This exon was part of a predicted cistron that is 24,195 bp in length. The sequence corresponding to the cistron was noted and saved separately from the 200,757 bp segment. BLASTN searches of the Unigene, dbEST, and nr databases were performed on it.

At least 100 significant alignments to various regions of the sequence were found in the dbEST database, although most appeared to be redundant representations of a few exons. All matches were to unnamed cDNAs and mRNAs (unnamed cDNAs and mRNAs are cDNAs and mRNAs not previously identified, or shown to correspond to a known or predicted human gene) from various tissue types. Most aligned to a single region on the sequence and spanned 500 bp or less, but several consisted of five or six regions separated by gaps, suggesting the locations of exons in the gene. Several significant matches to entries in the UniGene database were found, as well, even after masking low-complexity regions and short repeats in the sequence. All matches were to unnamed cDNA clones.

At least 100 significant alignments were found in the nr database, as well. A similarity to hypothetical protein FLJ22457 (UniGene cluster Hs.238707)was found (e=0.0). The cDNA of this predicted protein has been isolated from B lymphocytes

(http://www.ncbi.nlm.nih.gov/entrez/viewer.cgi?save=0&cmd=&cfm=on&f=1&view=gp&txt=0&val=13637988).

Other significant alignments were to unnamed cDNAs and mRNAs.

Using Genscan, the following 730 residue peptide sequence was predicted from the putative gene:

MDGLGRRLRA	SLRLKRGHGG	HWRLNEMPYM	KHEFDGGPPQ	DNSGEALKEP	5
ERAQEHSLPN	FAGGQHFFEY	LLVVSLKKKR	SEDDYEPIIT	YQFPKRENLL	1
RGQQEEEERL	LKAIPLFCFP	DGNEWASLTE	YPSLSCKTPG	LLAALVVEKA	1
QPRTCCHASA	PSAAPQARGP	DAPSPAAGQA	LPAGPGPRLP	KVYCIISCIG	2
CFGLFSKILD	EVEKRHQISM	AVIYPFMQGL	REAAFPAPGK	TVTLKSFIPD	2
SGTEFISLTR	PLDSHLEHVD	FSSLLHCLSF	EQILQIFASA	VLERKIIFLA	3
EGLREEEKDV	RDSTEVRGAG	ECHGFQRKGN	LGKQWGLCVE	DSVKMGDNQR	3
GTSCSTLSQC	IHAAAALLYP	FSWAHTYIPV	VPESLLATVC	CPTPFMVGVQ	4
MRFQQEVMDS	PMEEIQPQAE	IKTVNPLGVY	EERGPEKASL	CLFQVLLVNL	4
CEGTFLMSVG	DEKDILPPKL	QDDILDSLGQ	GINELKTAEQ	INEHVSGPFV	5
QFFVKIVGHY	ASYIKREANG	QGHFQERSFC	KALTSKTNRR	FVKKFVKTQL	5
FSLFIQEAEK	SKNPPAEVTQ	VGNSSTCVVD	TWLEAAATAL	SHHYNIFNTE	6
HTLWSKGSAS	LHEVCGHVRT	RVKRKILFLY	VSLAFTMGKS	IFLVENKAMN	6
MTIKWTTSGR	PGHGDMFGVI	ESWGAAALLL	LTGRVRDTGK	SSSSTGHRAS	7
KSLVWSQVCF	PESWEERLLT	EGKQLQSRVI	SEQ ID NO:8768		

Multiple analyses were performed using this prediction. First, a pairwise comparison of the sequence above and the sequence of FLJ22457, the hypothetical protein mentioned above, using BLASTP version 2.1.2 (http://ncbi.nlm.nih.gov/BLAST/), resulted in a match with an expect value of 0.0. The peptide sequence predicted from clone 596H6 was longer and 19% of the region of alignment between the two resulted from gaps in hypothetical protein FLJ22457. The cause of the discrepancy might be alternative mRNA splicing, alternative post-translational processing, or differences in the peptide-predicting algorithms used to create the two sequences, but the homology between the two is significant.

BLASTP and TBLASTN were also used to search for sequence similarities in the SWISS-PROT, TrEMBL, GenBank Translated, and PDB databases. Matches to several proteins were found, among them a tumor cell suppression protein, HTS1. No

matches aligned to the full length of the peptide sequence, however, suggesting that similarity is limited to a few regions of the peptide.

TBLASTN produced matches to several proteins – both identified and theoretical – but again, no matches aligned to the full length of the peptide sequence. The best alignment was to the same hypothetical protein found in GenBank before (FLJ22457).

To discover similarities to protein families, comparisons of the domains (described above) were carried out using the Pfam and Blocks databases. A search of the Pfam database identified two regions of the peptide domains as belonging the DENN protein family ($e=2.1 \times 10^{-33}$). The human DENN protein possesses an RGD cellular adhesion motif and a leucine-zipper-like motif associated with protein dimerization, and shows partial homology to the receptor binding domain of tumor necrosis factor alpha. DENN is virtually identical to MADD, a human MAP kinase-activating death domain protein that interacts with type I tumor necrosis factor receptor (http://srs.ebi.ac.uk/srs6bin/cgi-bin/wgetz?-id+fS5n1GQsHf+-e+[INTERPRO:'IPR001194']). The search of the Blocks database also revealed similarities between regions of the peptide sequence and known protein groups, but none with a satisfactory degree of confidence. In the Blocks scoring system, scores over 1,100 are likely to be relevant. The highest score of any match to the predicted peptide was 1,058.

The Prosite, ProDom, PRINTS databases (all publicly available) were used to conduct further domain and motif analysis. The Prosite search generated many recognized protein domains. A BLASTP search was performed to identify areas of similarity between the protein query sequence and PRINTS, a protein database of protein fingerprints, groups of motifs that together form a characteristic signature of a protein family. In this case, no groups were found to align closely to any section of the submitted sequence. The same was true when the ProDom database was searched with BLASTP.

A prediction of protein structure was done by performing a BLAST search of the sequence against PDB, a database in which every member has tertiary structure information. No significant alignments were found by this method. Secondary and super-secondary structure was examined using the Garnier algorithm. Although it is only considered to be 60-65% accurate, the algorithm provided information on the locations and lengths of alpha-helices, beta-sheets, turns and coils.

The antigenicity of the predicted peptide was modeled by graphing hydrophilicity vs. amino acid number. This produced a visual representation of trends in hydrophilicity along the sequence. Many locations in the sequence showed antigenicity and five sites had antigenicity greater than 2. This information can be used in the design of affinity reagents to the protein.

Membrane-spanning regions were predicted by graphing hydrophobicity vs. amino acid number. Thirteen regions were found to be somewhat hydrophobic. The algorithm TMpred predicted a model with 6 strong transmembrane helices (http://www.ch.embnet.org/software/

TMPRED_form.html).

NNPSL is a neural network algorithm developed by the Sanger Center. It uses amino acid composition and sequence to predict cellular location. For the peptide sequence submitted, its first choice was mitochondrial (51.1% expected accuracy). Its second choice was cytoplasmic (91.4% expected accuracy).

Example 6: Further sequence analysis of novel Clone 486E11

The sequence of clone 486E11 is provided below:

TAAAAGCAGG	CTGTGCACTA	GGGACCTAGT	GACCTTACTA	GAAAAAACTC	5
AAATTCTCTG	AGCCACAAGT	CCTCATGGGC	AAAATGTAGA	TACCACCACC	1
TAACCCTGCC	AATTTCCTAT	CATTGTGACT	ATCAAATTAA	ACCACAGGCA	1
GGAAGTTGCC	TTGAAAACTT	TTTATAGTGT	ATATTACTGT	TCACATAGAT	2
NAGCAATTAA	CTTTACATAT	ACCCGTTTTT	AAAAGATCAG	TCCTGTGATT	2
AAAAGTCTGG	CTGCCCTAAT	TCACTTCGAT	TATACATTAG	GTTAAAGCCA	3
TATAAAAGAG	GCACTACGTC	TTCGGAGAGA	TGAATGGATA	TTACAAGCAG	3
TAATGTTGGC	TTTGGAATAT	ACACATAATG	TCCACTTGAC	CTCATCTATT	4
TGACACAAAA	TGTAAACTAA	ATTATGAGCA	TCATTAGATA	CCTTGGCCTT	4
TTCAAATCAC	ACAGGGTCCT	AGATCTNNNN	NNNNNNNNNN	NNNNNNNNNN	5
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNAC	TTTGGGATTC	5
CTATATCTTT	GTCAGCTGTC	AACTTCAGTG	TTTTCAGGTT	AAATTCTATC	6
CATAGTCATC	CCAATATACC	TGCTTTAGAT	GATACAACCT	TCAAAAGATC	6
CGCTCTTCCT	CGTAAAAAGT	GGAG	SEQ ID NO: 8769		

The BLASTN program was used to compare the sequence to the UniGene and dbEST databases. No significant alignments were found in either. It was then searched against the nr database and only alignments to unnamed genomic DNA clones were found.

CAP2 was used to cluster a group of unknowns, including clone 486E11. The sequence for 486E11 was found to overlap others. These formed a contig of 1,010 residues, which is shown below:

CGGACAGGTA CCTAAAAGCA GGCTGTGCAC TAGGGACCTA GTGACCTTAC 5
TAGAAAAAAC TCAAATTCTC TGAGCCACAA GTCCTCATGG GCAAAATGTA 1
GATACCACCA CCTAACCCTG CCAATTTCCT ATCATTGTGA CTATCAAATT 1
AAACCACAGG CAGGAAGTTG CCTTGAAAAC TTTTTATAGT GTATATTACT 2
GTTCACATAG ATNAGCAATT AACTTTACAT ATACCCGTTT TTAAAAGATC 2
AGTCCTGTGA TTAAAAGTCT GGCTGCCCTA ATTCACTTCG ATTATACATT 3
AGGTTAAAGC CATATAAAAG AGGCACTACG TCTTCGGAGA GATGAATGGA 3
TATTACAAGC AGTAATTTTG GCTTTGGAAT ATACACATAA TGTCCACTTG 4
ACCTCATCTA TTTGACACAA AATGTAAACT AAATTATGAG CATCATTAGA 4
TACCTTGGGC CTTTTCAAAT CACACAGGGT CCTAGATCTG NNNNNNNNNN 5
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 5
NACTTTGGAT TCTTATATCT TTGTCAGCTG TCAACTTCAG TGTTTTCAGG 6
NTAAATTCTA TCCATAGTCA TCCCAATATA CCTGCTTTAG ATGATACAAA 6
CTTCAAAAGA TCCGGCTCTC CCTCGTAAAA CGTGGAGGAC AGACATCAAG 7
GGGGTTTTCT GAGTAAAGAA AGGCAACCGC TCGGCAAAAA CTCACCCTGG 7
CACAACAGGA NCGAATATAT ACAGACGCTG ATTGAGCGTT TTGCTCCATC 8
TTCACTTCTG TTAAATGAAG ACATTGATAT CTAAAATGCT ATGAGTCTAA 8
CTTTGTAAAA TTAAAATAGA TTTGTAGTTA TTTTTCAAAA TGAAATCGAA 9
AAGATACAAG TTTTGAAGGC AGTCTCTTTT TCCACCCTGC CCCTCTAGTG 9
TGTTTTACAC ACTTCTCTGG CCACTCCAAC AGGGAAGCTG GTCCAGGGCC 1
ATTATACAGG SEQ ID NO: 8832

The sequence of the CAP2 contig was used in a BLAST search of the human genome. 934 out of 1,010 residues aligned to a region of chromosome 21. A gap of 61 residues divided the aligned region into two smaller fragments. The sequence of this region, plus 100 kilobases on each side of it, was downloaded and analyzed using the Genscan site at MIT (http://genes.mit.edu/GENSCAN.html), with the following settings:

Organism: vertebrate

Suboptimal exon cutoff: 1.00 (no suboptimal exons)

Print options: Predicted CDS and peptides

The fragment was found to fall within one of several predicted genes in the chromosome region. The bases corresponding to the predicted gene, including its predicted introns, were saved as a separate file and used to search GenBank again with BLASTN to find any ESTs or UniGene clusters identified by portions of the sequence not included in the original unknown fragment. The nr database contained no significant matches. At least 100 significant matches to various parts of the predicted gene were found in the dbEST database, but all of them were to unnamed cDNA clones. Comparison to UniGene produced fewer significant matches, but all matches were to unnamed cDNAs.

The peptide sequence predicted by Genscan was also saved. Multiple types of analyses were performed on it using the resources mentioned in Table 3. BLASTP and TBLASTN were used to search the TrEMBL protein database (http://www.expasy.ch/sprot/) and the GenBank nr database (http://www.ncbi.nlm.hih.gov/BLAST/), which includes data from the SwissProt, PIR, PRF, and PDB databases. No significant matches were found in any of these, so no gene identity or tertiary structure was discovered.

The peptide sequence was also searched for similarity to known domains and motifs using BLASTP with the Prosite, Blocks, Pfam, and ProDom databases. The searches produced no significant alignments to known domains. BLASTP comparison to the PRINTS database produced an alignment to the P450 protein family, but with a low probability of accuracy (e=6.9).

Two methods were used to predict secondary structure – the Garnier/Osguthorpe/Robson model and the Chou-Fasman model. The two methods differed somewhat in their results, but both produced representations of the peptide sequence with helical and sheet regions and locations of turns.

Antigenicity was plotted as a graph with amino acid number in the sequence on the x-axis and hydrophilicity on the y-axis. Several areas of antigenicity were observed, but only one with antigenicity greater than 2. Hydrophobicity was plotted in the same way. Only one region, from approximately residue 135 to residue 150, had notable hydrophobicity. TMpred, accessed through ExPASy, was used to predict transmembrane helices. No regions of the peptide sequence were predicted with reasonable confidence to be membrane-spanning helices.

NNPSL predicted that the putative protein would be found either in the nucleus (expected prediction accuracy = 51.1%) or secreted from the cell (expected prediction accuracy = 91.4%).

Example 7: Preparation of a leukocyte cDNA array comprising a candidate gene library

Candidate genes and gene sequences for leukocyte expression profiling were identified through methods described elsewhere in this document. Candidate genes are used to obtain or design probes for peripheral leukocyte expression profiling in a variety of ways.

A cDNA microarray carrying 384 probes was constructed using sequences selected from the cDNA libraries described in example 1. cDNAs were selected from T-cell libraries, PBMC libraries and buffy coat libraries. A listing of the cDNA fragments used is given in Table 8.

96-Well PCR

Plasmids were isolated in 96-well format and PCR was performed in 96-well format. A master mix was made that contain the reaction buffer, dNTPs, forward and reverse primer and DNA polymerase was made. 99 ul of the master mix was aliquoted into 96-well plate. 1 ul of plasmid (1-2 ng/ul) of plasmid was added to the plate. The final reaction concentration was 10 mM Tris pH 8.3, 3.5 mM MgCl2, 25 mM KCl, 0.4 mM dNTPs, 0.4 uM M13 forward primer, 0.4 M13 reverse primer, and 10 U of Taq Gold (Applied Biosystems). The PCR conditions were:

Step 1 95C for 10 min

Step 2 95C for 15 sec

Step 3 56C for 30 sec

Step 4 72C for 2 min 15 seconds

Step 5 go to Step 2 39 times

Step 6 72C for 10 minutes

Step 7 4C for ever.

PCR Purification

PCR purification was done in a 96-well format. The ArrayIt (Telechem International, Inc.) PCR purification kit was used and the provided protocol was followed without modification. Before the sample was evaporated to dryness, the

concentration of PCR products was determined using a spectrophotometer. After evaporation, the samples were re-suspended in 1x Micro Spotting Solution (ArrayIt) so that the majority of the samples were between 0.2-1.0 ug/ul.

Array Fabrication

Spotted cDNA microarrays were then made from these PCR products by ArrayIt using their protocols (http://arrayit.com/Custom_Microarrays/Flex-Chips/flex-chips.html). Each fragment was spotted 3 times onto each array.

Candidate genes and gene sequences for leukocyte expression profiling were identified through methods described elsewhere in this document. Those candidate genes are used for peripheral leukocyte expression profiling. The candidate libraries can used to obtain or design probes for expression profiling in a variety of ways.

Oligonucleotide probes are also prepared using the DNA sequence information for the candidate genes identified by differential hybridization screening (listed in Table 3 and the sequence listing) and/or the sequence information for the genes identified by database mining (listed in Table 2) is used to design complimentary oligonucleotide probes. Oligo probes are designed on a contract basis by various companies (for example, Compugen, Mergen, Affymetrix, Telechem), or designed from the candidate sequences using a variety of parameters and algorithms as indicated at http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi. Briefly, the length of the oligonucleotide to be synthesized is determined, preferably greater than 18 nucleotides, generally 18-24 nucleotides, 24-70 nucleotides and, in some circumstances, more than 70 nucleotides. The sequence analysis algorithms and tools described above are applied to the sequences to mask repetitive elements, vector sequences and low complexity sequences. Oligonucleotides are selected that are specific to the candidate nucleotide sequence (based on a Blast n search of the oligonucleotide sequence in question against gene sequences databases, such as the Human Genome Sequence, UniGene, dbEST or the non-redundant database at NCBI), and have <50% G content and 25–70% G+C content. Desired oligonucleotides are synthesized using well-known methods and apparatus, or ordered from a company (for example Sigma). Oligonucleotides are spotted onto microarrays. Alternatively, oligonucleotides are synthesized directly on the array surface, using a variety of techniques (Hughes et al. 2001, Yershov et al. 1996, Lockhart et al 1996).

Example 8: Preparation of RNA from mononuclear cells for expression profiling

Blood was isolated from the subject for leukocyte expression profiling using the following methods:

Two tubes were drawn per patient. Blood was drawn from either a standard peripheral venous blood draw or directly from a large-bore intra-arterial or intravenous catheter inserted in the femoral artery, femoral vein, subclavian vein or internal jugular vein. Care was taken to avoid sample contamination with heparin from the intravascular catheters, as heparin can interfere with subsequent RNA reactions.

For each tube, 8 ml of whole blood was drawn into a tube (CPT, Becton-Dickinson order #362753) containing the anticoagulant Citrate, 25°C density gradient solution (e.g. Ficoll, Percoll) and a polyester gel barrier that upon centrifugation was permeable to RBCs and granulocytes but not to mononuclear cells. The tube was inverted several times to mix the blood with the anticoagulant. The tubes were centrifuged at 1750xg in a swing-out rotor at room temperature for 20 minutes. The tubes were removed from the centrifuge and inverted 5-10 times to mix the plasma with the mononuclear cells, while trapping the RBCs and the granulocytes beneath the gel barrier. The plasma/mononuclear cell mix was decanted into a 15ml tube and 5ml of phosphate-buffered saline (PBS) is added. The 15ml tubes were spun for 5 minutes at 1750xg to pellet the cells. The supernatant was discarded and 1.8 ml of RLT lysis buffer is added to the mononuclear cell pellet. The buffer and cells were pipetted up and down to ensure complete lysis of the pellet. The cell lysate was frozen and stored until it is convenient to proceed with isolation of total RNA.

Total RNA was purified from the lysed mononuclear cells using the Qiagen Rneasy Miniprep kit, as directed by the manufacturer (10/99 version) for total RNA isolation, including homogenization (Qiashredder columns) and on-column DNase treatment. The purified RNA was eluted in 50ul of water. The further use of RNA prepared by this method is described in Example 11, 24, and 23.

Some samples were prepared by a different protocol, as follows:

Two 8 ml blood samples were drawn from a peripheral vein into a tube (CPT, Becton-Dickinson order #362753) containing anticoagulant (Citrate), 25°C density gradient solution (Ficoll) and a polyester gel barrier that upon centrifugation is permeable to RBCs and granulocytes but not to mononuclear cells. The mononuclear cells and plasma remained above the barrier while the RBCs and granulocytes were

trapped below. The tube was inverted several times to mix the blood with the anticoagulant, and the tubes were subjected to centrifugation at 1750xg in a swing-out rotor at room temperature for 20 min. The tubes were removed from the centrifuge, and the clear plasma layer above the cloudy mononuclear cell layer was aspirated and discarded. The cloudy mononuclear cell layer was aspirated, with care taken to rinse all of the mononuclear cells from the surface of the gel barrier with PBS (phosphate buffered saline). Approximately 2 mls of mononuclear cell suspension was transferred to a 2ml microcentrifuge tube, and centrifuged for 3min. at 16,000 rpm in a microcentrifuge to pellet the cells. The supernatant was discarded and 1.8 ml of RLT lysis buffer (Qiagen) were added to the mononuclear cell pellet, which lysed the cells and inactivated Rnases. The cells and lysis buffer were pipetted up and down to ensure complete lysis of the pellet. Cell lysate was frozen and stored until it was convenient to proceed with isolation of total RNA.

RNA samples were isolated from 8 mL of whole blood. Yields ranged from 2 ug to 20ug total RNA for 8mL blood. A260/A280 spectrophotometric ratios were between 1.6 and 2.0, indicating purity of sample. 2ul of each sample were run on an agarose gel in the presence of ethidium bromide. No degradation of the RNA sample and no DNA contamination was visible.

Example 9: Preparation of Buffy Coat Control RNA for use in leukocyte expression profiling

Control RNA was prepared using total RNA from Buffy coats and/or total RNA from enriched mononuclear cells isolated from Buffy coats, both with and without stimulation with ionomycin and PMA. The following control RNAs were prepared:

Control 1: Buffy Coat Total RNA

Control 2: Mononuclear cell Total RNA

Control 3: Stimulated buffy coat Total RNA

Control 4: Stimulated mononuclear Total RNA

Control 5: 50% Buffy coat Total RNA / 50% Stimulated buffy coat Total RNA

Control 6: 50% Mononuclear cell Total RNA / 50% Stimulated Mononuclear Total RNA

Some samples were prepared using the following protocol: Buffy coats from 38 individuals were obtained from Stanford Blood Center. Each buffy coat is derived from ~350 mL whole blood from one individual. 10 ml buffy coat was removed from the bag, and placed into a 50 ml tube. 40 ml of Buffer EL (Qiagen) was added, the tube was mixed and placed on ice for 15 minutes, then cells were pelleted by centrifugation at 2000xg for 10 minutes at 4°C. The supernatant was decanted and the cell pellet was re-suspended in 10 ml of Qiagen Buffer EL. The tube was then centrifuged at 2000xg for 10 minutes at 4°C. The cell pellet was then re-suspended in 20 ml TRIZOL (GibcoBRL) per Buffy coat sample, the mixture was shredded using a rotary homogenizer, and the lysate was then frozen at –80°C prior to proceeding to RNA isolation.

Other control RNAs were prepared from enriched mononuclear cells prepared from Buffy coats. Buffy coats from Stanford Blood Center were obtained, as described above. 10 ml buffy coat was added to a 50 ml polypropylene tube, and 10 ml of phosphate buffer saline (PBS) was added to each tube. A polysucrose (5.7 g/dL) and sodium diatrizoate (9.0 g/dL) solution at a 1.077 +/-0.0001 g/ml density solution of equal volume to diluted sample was prepared (Histopaque 1077, Sigma cat. no 1077-1). This and all subsequent steps were performed at room temperature. 15 ml of diluted buffy coat/PBS was layered on top of 15 ml of the histopaque solution in a 50 ml tube. The tube was centrifuged at 400xg for 30 minutes at room temperature. After centrifugation, the upper layer of the solution to within 0.5 cm of the opaque interface containing the mononuclear cells was discarded. The opaque interface was transferred into a clean centrifuge tube. An equal volume of PBS was added to each tube and centrifuged at 350xg for 10 minutes at room temperature. The supernatant was discarded. 5 ml of Buffer EL (Qiagen) was used to resuspend the remaining cell pellet and the tube was centrifuged at 2000xg for 10 minutes at room temperature. The supernatant was discarded. The pellet was resuspended in 20 ml of TRIZOL (GibcoBRL) for each individual buffy coat that was processed. The sample was homogenized using a rotary homogenizer and frozen at -80C until RNA was isolated.

RNA was isolated from frozen lysed Buffy coat samples as follows: frozen samples were thawed, and 4 ml of chloroform was added to each buffy coat sample. The sample was mixed by vortexing and centrifuged at 2000xg for 5 minutes. The aqueous layer was moved to new tube and then repurified by using the RNeasy Maxi

RNA clean up kit, according to the manufacturer's instruction (Qiagen, PN 75162). The yield, purity and integrity were assessed by spectrophotometer and gel electrophoresis.

Some samples were prepared by a different protocol, as follows. The further use of RNA prepared using this protocol is described in Example 11.

50 whole blood samples were randomly selected from consented blood donors at the Stanford Medical School Blood Center. Each buffy coat sample was produced from ~350 mL of an individual's donated blood. The whole blood sample was centrifuged at ~4,400 x g for 8 minutes at room temperature, resulting in three distinct layers: a top layer of plasma, a second layer of buffy coat, and a third layer of red blood cells. 25 ml of the buffy coat fraction was obtained and diluted with an equal volume of PBS (phosphate buffered saline). 30 ml of diluted buffy coat was layered onto 15 ml of sodium diatrizoate solution adjusted to a density of 1.077+/-0.001 g/ml (Histopaque 1077, Sigma) in a 50mL plastic tube. The tube was spun at 800 g for 10 minutes at room temperature. The plasma layer was removed to the 30 ml mark on the tube, and the mononuclear cell layer removed into a new tube and washed with an equal volume of PBS, and collected by centrifugation at 2000 g for 10 minutes at room temperature. The cell pellet was resuspended in 10 ml of Buffer EL (Qiagen) by vortexing and incubated on ice for 10 minutes to remove any remaining erthythrocytes. The mononuclear cells were spun at 2000 g for 10 minutes at 4 degrees Celsius. The cell pellet was lysed in 25 ml of a phenol/guanidinium thiocyanate solution (TRIZOL Reagent, Invitrogen). The sample was homogenized using a PowerGene 5 rotary homogenizer (Fisher Scientific) and Omini disposable generator probes (Fisher Scientific). The Trizol lysate was frozen at –80 degrees C until the next step.

The samples were thawed out and incubated at room temperature for 5 minutes. 5 ml chloroform was added to each sample, mixed by vortexing, and incubated at room temperature for 3 minutes. The aqueous layers were transferred to new 50 ml tubes. The aqueous layer containing total RNA was further purified using the Qiagen RNeasy Maxi kit (PN 75162), per the manufacturer's protocol (October 1999). The columns were eluted twice with 1 ml Rnase-free water, with a minute incubation before each spin. Quantity and quality of RNA was assessed using standard methods. Generally, RNA was isolated from batches of 10 buffy coats at a

time, with an average yield per buffy coat of 870 μg, and an estimated total yield of 43.5 mg total RNA with a 260/280 ratio of 1.56 and a 28S/18S ratio of 1.78.

Quality of the RNA was tested using the Agilent 2100 Bioanalyzer using RNA 6000 microfluidics chips. Analysis of the electrophorgrams from the Bioanalyzer for five different batches demonstrated the reproducibility in quality between the batches.

Total RNA from all five batches were combined and mixed in a 50 ml tube, then aliquoted as follows: 2 x 10 ml aliquots in 15 ml tubes, and the rest in 100 μl aliquots in 1.5 ml microcentrifuge tubes. The aliquots gave highly reproducible results with respect to RNA purity, size and integrity. The RNA was stored at –80°C.

Test hybridization of Reference RNA

The reference RNA (hereinafter, "R50") was hybridized to a spotted cDNA array (prepared as described in Example 10). There are a total of 1152 features on the array: 384 clones printed in triplicate. The R50 targets were fluorescently labeled with Cy-5 using methods described herein. In five array hybridizations, the reference RNA detected 94% of probes on the array with a Signal to Noise ratio of greater than three. 99% of probes on the array were detected with a signal to noise ratio of greater than one. Figure 8 shows one array hybridization. The probes are ordered from high to low in signal to noise ratio, and the log of median and the log of the background were plotted for each probe.

Example 10. RNA Labeling and hybridization to a leukocyte cDNA array of candidate nucleotide sequences.

Comparison of Guanine-Silica to Acid-Phenol RNA Purification (GSvsAP)

These data are from a set of 12 hybridizations designed to identify differences between the signal strength from two different RNA purification methods. The two RNA methods used were guanidine-silica (GS, Qiagen) and acid-phenol (AP, Trizol, Gibco BRL). Ten tubes of blood were drawn from each of four people. Two were used for the AP prep, the other eight were used for the GS prep. The protocols for the leukocyte RNA preps using the AP and GS techniques were completed as described here:

Guanidine-silica (GS) method:

For each tube, 8ml blood was drawn into a tube containing the anticoagulant Citrate, 25°C density gradient solution and a polyester gel barrier that upon centrifugation is permeable to RBCs and granulocytes but not to mononuclear cells.

The mononuclear cells and plasma remained above the barrier while the RBCs and granulocytes were trapped below. CPT tubes from Becton-Dickinson (#362753) were used for this purpose. The tube was inverted several times to mix the blood with the anticoagulant. The tubes were immediately centrifuged @1750xg in a swinging bucket rotor at room temperature for 20 min. The tubes were removed from the centrifuge and inverted 5-10 times. This mixed the plasma with the mononuclear cells, while the RBCs and the granulocytes remained trapped beneath the gel barrier. The plasma/mononuclear cell mix was decanted into a 15ml tube and 5ml of phosphate-buffered saline (PBS) was added. The 15ml tubes are spun for 5 minutes at 1750xg to pellet the cells. The supernatant was discarded and 1.8 ml of RLT lysis buffer (guanidine isothyocyanate) was added to the mononuclear cell pellet. The buffer and cells were pipetted up and down to ensure complete lysis of the pellet. The cell lysate was then processed exactly as described in the Qiagen Rneasy Miniprep kit protocol (10/99 version) for total RNA isolation (including steps for homogenization (Qiashredder columns) and on-column DNase treatment. The purified RNA was eluted in 50ul of water.

Acid-phenol (AP) method:

For each tube, 8ml blood was drawn into a tube containing the anticoagulant Citrate, 25°C density gradient solution and a polyester gel barrier that upon centrifugation is permeable to RBCs and granulocytes but not to mononuclear cells. The mononuclear cells and plasma remained above the barrier while the RBCs and granulocytes were trapped below. CPT tubes from Becton-Dickinson (#362753) were used for this purpose. The tube was inverted several times to mix the blood with the anticoagulant. The tubes were immediately centrifuged @1750xg in a swinging bucket rotor at room temperature for 20 min. The tubes were removed from the centrifuge and inverted 5-10 times. This mixed the plasma with the mononuclear cells, while the RBCs and the granulocytes remained trapped beneath the gel barrier. The plasma/mononuclear cell mix was decanted into a 15ml tube and 5ml of phosphate-buffered saline (PBS) was added. The 15ml tubes are spun for 5 minutes @1750xg to pellet the cells. The supernatant was discarded and the cell pellet was lysed using 0.6 mL Phenol/guanidine isothyocyanate (e.g. Trizol reagent, GibcoBRL). Subsequent total RNA isolation proceeded using the manufacturers protocol.

RNA from each person was labeled with either Cy3 or Cy5, and then hybridized in pairs to the mini-array. For instance, the first array was hybridized with GS RNA from one person (Cy3) and GS RNA from a second person (Cy5).

Techniques for labeling and hybridization for all experiments discussed here were completed as detailed above in example 10. Arrays were prepared as described in example 7.

RNA isolated from subject samples, or control Buffy coat RNA, were labeled for hybridization to a cDNA array. Total RNA (up to 100 μg) was combined with 2 μl of 100 μM solution of an Oligo (dT)12-18 (GibcoBRL) and heated to 70°C for 10 minutes and place on ice. Reaction buffer was added to the tube, to a final concentration of 1xRT buffer (GibcoBRL), 10 mM DTT (GibcoBRL), 0.1 mM unlabeled dATP, dTTP, and dGTP, and 0.025 mM unlabeled dCTP, 200 pg of CAB (A. thaliana photosystem I chlorophyll a/b binding protein), 200 pg of RCA (A. thaliana RUBISCO activase), 0.25 mM of Cy-3 or Cy-5 dCTP, and 400 U Superscript II RT (GibcoBRL).

The volumes of each component of the labeling reaction were as follows: 20 μl of 5xRT buffer; 10 μl of 100 mM DTT; 1 μl of 10 mM dNTPs without dCTP; 0.5 μl of 5 mM CTP; 13 μl of H20; 0.02 μl of 10 ng/μl CAB and RCA; 1 μl of 40 Units/μl RNAseOUT Recombinatnt Ribonuclease Inhibitor (GibcoBRL); 2.5 μl of 1.0 mM Cy-3 or Cy-5 dCTP; and 2.0 μl of 200 Units/μl of Superscript II RT. The sample was vortexed and centrifuged. The sample was incubated at 4°C for 1 hour for first strand cDNA synthesis, then heated at 70°C for 10 minutes to quench enzymatic activity. 1 μl of 10 mg/ml of Rnase A was added to degrade the RNA strand, and the sample was incubated at 37°C for 30 minutes.

Next, the Cy-3 and Cy-5 cDNA samples were combined into one tube. Unincorporated nucleotides were removed using QIAquick RCR purification protocol (Qiagen), as directed by the manufacturer. The sample was evaporated to dryness and resuspended in 5 μl of water. The sample was mixed with hybridization buffer containing 5xSSC, 0.2% SDS, 2 mg/ml Cot-1 DNA (GibcoBRL), 1 mg/ml yeast tRNA (GibcoBRL), and 1.6 ng/μl poly dA40-60 (Pharmacia). This mixture was placed on the microarray surface and a glass cover slip was placed on the array (Corning). The microarray glass slide was placed into a hybridization chamber (ArrrayIt). The chamber was then submerged in a water bath overnight at 62° C. The

microarray was removed from the cassette and the cover slip was removed by repeatedly submerging it to a wash buffer containing 1xSSC, and 0.1% SDS. The microarray slide was washed in 1xSSC/0.1% SDS for 5 minutes. The slide was then washed in 0.1%SSC/0.1% SDS for 5 minutes. The slide was finally washed in 0.1xSSC for 2 minutes. The slide was spun at 1000 rpm for 2 minutes to dry out the slide, then scanned on a microarray scanner (Axon Instruments, Union City, CA.).

Six hybridizations with 20 μg of RNA were performed for each type of RNA preparation (GS or AP). Since both the Cy3 and the Cy5 labeled RNA are from test preparations, there are six data points for each GS prepped, Cy3-labeled RNA and six for each GS-prepped, Cy5-labeled RNA. The mini array hybridizations were scanned on and Axon Instruments scanner using GenPix 3.0 software. The data presented were derived as follows. First, all features flagged as "not found" by the software were removed from the dataset for individual hybridizations. These features are usually due to high local background or other processing artifacts. Second, the median fluorescence intensity minus the background fluorescence intensity was used to calculate the mean background subtracted signal for each dye for each hybridization. In Figure 4, the mean of these means across all six hybridizations is graphed (n=6 for each column). The error bars are the SEM. This experiment shows that the average signal from AP prepared RNA is 47% of the average signal from GS prepared RNA for both Cy3 and Cy5.

Generation of expression data for leukocyte genes from peripheral leukocyte samples

Six hybridizations were performed with RNA purified from human blood leukocytes using the protocols given above. Four of the six were prepared using the GS method and 2 were prepared using the AP method. Each preparation of leukocyte RNA was labeled with Cy3 and 10 μg hybridized to the mini-array. A control RNA was batch labeled with Cy5 and 10 μg hybridized to each mini-array together with the Cy3-labeled experimental RNA.

The control RNA used for these experiments was Control 1: Buffy Coat RNA, as described above. The protocol for the preparation of that RNA is reproduced here:

Buffy Coat RNA Isolation:

Buffy coats were obtained from Stanford Blood Center (in total 38 individual buffy coats were used. Each buffy coat is derived from ~350 mL whole blood from

one individual. 10 ml buffy coat was taken and placed into a 50 ml tube and 40 ml of a hypoclorous acid (HOCl) solution (Buffer EL from Qiagen) was added. The tube was mixed and placed on ice for 15 minutes. The tube was then centrifuged at 2000xg for 10 minutes at 4°C. The supernatant was decanted and the cell pellet was re-suspended in 10 ml of hypochlorous acid solution (Qiagen Buffer EL). The tube was then centrifuged at 2000xg for 10 minutes at 4°C. The cell pellet was then re-suspended in 20 ml phenol/guanidine thiocyanate solution (TRIZOL from GibcoBRL) for each individual buffy coat that was processed. The mixture was then shredded using a rotary homogenizer. The lysate was then frozen at –80°C prior to proceeding to RNA isolation.

The arrays were then scanned and analyzed on an Axon Instruments scanner using GenePix 3.0 software. The data presented were derived as follows. First, all features flagged as "not found" by the software were removed from the dataset for individual hybridizations. Second, control features were used to normalize the data for labeling and hybridization variability within the experiment. The control features are cDNA for genes from the plant, *Arabidopsis thaliana*, that were included when spotting the mini-array. Equal amounts of RNA complementary to two of these cDNAs were added to each of the samples before they were labeled. A third was pre-labeled and equal amounts were added to each hybridization solution before hybridization. Using the signal from these genes, we derived a normalization constant (L_j) according to the following formula:

$$L_j = \frac{\frac{\sum_{i=1}^{N} BGSS_{j,i}}{N}}{\frac{\sum_{j=1}^{K} \frac{\sum_{i=1}^{N} BGSS_{j,i}}{N}}{K}}$$

where $BGSS_i$ is the signal for a specific feature as identified in the GenePix software as the median background subtracted signal for that feature, N is the number of *A. thaliana* control features, K is the number of hybridizations, and L is the normalization constant for each individual hybridization.

Using the formula above, the mean over all control features of a particular hybridization and dye (eg Cy3) was calculated. Then these control feature means for all Cy3 hybridizations were averaged. The control feature mean in one hybridization divided by the average of all hybridizations gives a normalization constant for that particular Cy3 hybridization.

The same normalization steps were performed for Cy3 and Cy5 values, both fluorescence and background. Once normalized, the background Cy3 fluorescence was subtracted from the Cy3 fluorescence for each feature. Values less than 100 were eliminated from further calculations since low values caused spurious results.

Figure 5 shows the average background subtracted signal for each of nine leukocyte-specific genes on the mini array. This average is for 3-6 of the above-described hybridizations for each gene. The error bars are the SEM. Figure 3: The ratio of Cy3 to Cy5 signal is shown for a number of genes. This ratio corrects for variability among hybridizations and allows comparison between experiments done at different times. The ratio is calculated as the Cy3 background subtracted signal divided by the Cy5 background subtracted signal. Each bar is the average for 3-6 hybridizations. The error bars are SEM.

Together, these results show that we can measure expression levels for genes that are expressed specifically in sub-populations of leukocytes. These expression measurements were made with only 10 μg of leukocyte total RNA that was labeled directly by reverse transcription. The signal strength can be increased by improved labeling techniques that amplify either the starting RNA or the signal fluorescence. In addition, scanning techniques with higher sensitivity can be used.

Genes in Figures 5 and 6:

Gene Name/Description	GenBank Accession Number	Gene Name Abbreviation
T cell-specific tyrosine kinase Mrna	L10717	TKTCS
Interleukin 1 alpha (IL 1) mRNA, complete cds	NM_000575	IL1A
T-cell surface antigen CD2 (T11) mRNA, complete cds	M14362	CD2
Interleukin-13 (IL-13) precursor gene, complete cds	U31120	IL-13
Thymocyte antigen CD1a mRNA, complete cds	M28825	CD1a

CD6 mRNA for T cell glycoprotein CDS	NM_006725	CD6
MHC class II HLA-DQA1 mRNA, complete cds	U77589	HLA-DQA1
Granulocyte colony-stimulating factor	M28170	CD19
Homo sapiens CD69 antigen	NM_001781	CD69

Example 11: Identification of diagnostic gene sets useful in diagnosis and treatment of Cardiac allograft rejection

An observational study was conducted in which a prospective cohort of cardiac transplant recipients were analyzed for associations between clinical events or rejection grades and expression of a leukocyte candidate nucleotide sequence library. Patients were identified at 4 cardiac transplantation centers while on the transplant waiting list or during their routing post-transplant care. All adult cardiac transplant recipients (new or re-transplants) who received an organ at the study center during the study period or within 3 months of the start of the study period were eligible. The first year after transplantation is the time when most acute rejection occurs and it is thus important to study patients during this period. Patients provided informed consent prior to study procedures.

Peripheral blood leukocyte samples were obtained from all patients at the following time points: prior to transplant surgery (when able), the same day as routinely scheduled screening biopsies, upon evaluation for suspected acute rejection (urgent biopsies), on hospitalization for an acute complication of transplantation or immunosuppression, and when Cytomegalovirus (CMV) infection was suspected or confirmed. Samples were obtained through a standard peripheral vein blood draw or through a catheter placed for patient care (for example, a central venous catheter placed for endocardial biopsy). When blood was drawn from a intravenous line, care was taken to avoid obtaining heparin with the sample as it can interfere with downstream reactions involving the RNA. Mononuclear cells were prepared from whole blood samples as described in Example 8. Samples were processed within 2 hours of the blood draw and DNA and serum were saved in addition to RNA. Samples were stored at –70° C or on dry ice and sent to the site of RNA preparation in a sealed container with ample dry ice. RNA was isolated from subject samples as

described in Example 8 and hybridized to a candidate library of differentially expressed leukocyte nucleotide sequences, as further described in Examples 20-22. Methods used for amplification, labeling, hybridization and scanning are described in example 23. Analysis of human transplant patient mononuclear cell RNA hybridized to a microarray is shown in Example 24.

From each patient, clinical information was obtained at the following time points: prior to transplant surgery (when available), the same day as routinely scheduled screening biopsies, upon evaluation for suspected acute rejection (e.g., urgent biopsies), on hospitalization for an acute complication of transplantation or immunosuppression, and when Cytomegalovirus (CMV) infection was suspected or confirmed. Data was collected directly from the patient, from the patient's medical record, from diagnostic test reports or from computerized hospital databases. It was important to collect all information pertaining to the study clinical correlates (diagnoses and patient events and states to which expression data is correlated) and confounding variables (diagnoses and patient events and states that may result in altered leukocyte gene expression. Examples of clinical data collected are: patient sex, date of birth, date of transplant, race, requirement for prospective cross match, occurrence of pre-transplant diagnoses and complications, indication for transplantation, severity and type of heart disease, history of left ventricular assist devices, all known medical diagnoses, blood type, HLA type, viral serologies (including CMV, Hepatitis B and C, HIV and others), serum chemistries, white and red blood cell counts and differentials, CMV infections (clinical manifestations and methods of diagnosis), occurrence of new cancer, hemodynamic parameters measured by catheterization of the right or left heart (measures of graft function), results of echocardiography, results of coronary angiograms, results of intravascular ultrasound studies (diagnosis of transplant vasculopathy), medications, changes in medications, treatments for rejection, and medication levels. Information was also collected regarding the organ donor, including demographics, blood type, HLA type, results of screening cultures, results of viral serologies, primary cause of brain death, the need for inotropic support, and the organ cold ischemia time.

Of great importance was the collection of the results of endocardial biopsy for each of the patients at each visit. Biopsy results were all interpreted and recorded using the international society for heart and lung transplantation (ISHLT) criteria, described below. Biopsy pathological grades were determined by experienced

pathologists at each center. It is desirable to have a single centralized pathologist determine the grades when an analysis is done using samples from multiple medical centers.

ISHLT Criteria

Grade	Finding	Rejection Severity
0	No lymphocytic infiltrates	None
1A	Focal (perivascular or interstitial lymphocytic infiltrates without necrosis)	Borderline mild
1B	Diffuse but sparse lymphocytic infiltrates without necrosis	Mild
2	One focus only with aggressive lymphocytic infiltrate and/or myocyte damage	Mild, focal moderate
3A	Multifocal aggressive lymphocytic infiltrates and/or myocardial damage	Moderate
3B	Diffuse inflammatory lymphocytic infiltrates with necrosis	Borderline Severe
4	Diffuse aggressive polymorphous lymphocytic infiltrates with edema hemorrhage and vasculitis, with necrosis	Severe

Clinical data was entered and stored in a database. The database was queried to identify all patients and patient visits that meet desired criteria (for example, patients with > grade II biopsy results, no CMV infection and time since transplant < 12 weeks).

The collected clinical data (disease criteria) is used to define patient or sample groups for correlation of expression data. Patient groups are identified for comparison, for example, a patient group that possesses a useful or interesting clinical distinction, versus a patient group that does not possess the distinction. Examples of useful and interesting patient distinctions that can be made on the basis of collected clinical data are listed here (and further described in Table 2):

1. Rejection episode of at least moderate histologic grade, which results in treatment of the patient with additional corticosteroids, anti-T cell antibodies, or total lymphoid irradiation.

2. Rejection with histologic grade 2 or higher.
3. Rejection with histologic grade <2.
4. The absence of histologic rejection and normal or unchanged allograft function (based on hemodynamic measurements from catheterization or on echocardiographic data).
5. The presence of severe allograft dysfunction or worsening allograft dysfunction during the study period (based on hemodynamic measurements from catheterization or on echocardiographic data).
6. Documented CMV infection by culture, histology, or PCR, and at least one clinical sign or symptom of infection.
7. Specific graft biopsy rejection grades
8. Rejection of mild to moderate histologic severity prompting augmentation of the patient's chronic immunosuppressive regimen
9. Rejection of mild to moderate severity with allograft dysfunction prompting plasmaphoresis or a diagnosis of "humoral" rejection
10. Infections other than CMV, esp. Epstein Barr virus (EBV)
11. Lymphoproliferative disorder (also called, post-transplant lymphoma)
12. Transplant vasculopathy diagnosed by increased intimal thickness on intravascular ultrasound (IVUS), angiography, or acute myocardial infarction.
13. Graft Failure or Retransplantation
14. All cause mortality

Expression profiles of subject samples are examined to discover sets of nucleotide sequences with differential expression between patient groups, for example, by methods describes above and below.

Non-limiting examples of patient leukocyte samples to obtain for discovery of various diagnostic nucleotide sets are as follows:

a. Leukocyte set to avoid biopsy or select for biopsy:
Samples : Grade 0 vs. Grades 1-4

b. Leukocyte set to monitor therapeutic response:
Examine successful vs. unsuccessful drug treatment.

Samples:
Successful: Time 1: rejection, Time 2: drug therapy Time 3: no rejection
Unsuccessful: Time 1: rejection, Time 2: drug therapy; Time 3: rejection

c. Leukocyte set to predict subsequent acute rejection.
Biopsy may show no rejection, but the patient may develop rejection shortly thereafter. Look at profiles of patients who subsequently do and do not develop rejection.

Samples:
Group 1 (Subsequent rejection): Time 1: Grade 0; Time 2: Grade>0

Group 2 (No subsequent rejection): Time 1: Grade 0, ; Time 2: Grade 0

Focal rejection may be missed by biopsy. When this occurs the patient may have a Grade 0, but actually has rejection. These patients may go on to have damage to the graft etc.

Samples:
Non-rejectors: no rejection over some period of time
Rejectors: an episode of rejection over same period

d. Leukocyte set to diagnose subsequent or current graft failure:

Samples:
Echocardiographic or catheterization data to define worsening function over time and correlate to profiles.

e. Leukocyte set to diagnose impending active CMV:

Samples:
Look at patients who are CMV IgG positive. Compare patients with subsequent (to a sample) clinical CMV infection verses no subsequent clinical CMV infection.

f. Leukocyte set to diagnose current active CMV:

Samples:
Analyze patients who are CMV IgG positive. Compare patients with active current clinical CMV infection vs. no active current CMV infection.

Upon identification of a nucleotide sequence or set of nucleotide sequences that distinguish patient groups with a high degree of accuracy, that nucleotide sequence or set of nucleotide sequences is validated, and implemented as a diagnostic test. The use of the test depends on the patient groups that are used to discover the nucleotide set. For example, if a set of nucleotide sequences is discovered that have collective expression behavior that reliably distinguishes patients with no histological rejection or graft dysfunction from all others, a diagnostic is developed that is used to screen patients for the need for biopsy. Patients identified as having no rejection do not need biopsy, while others are subjected to a biopsy to further define the extent of disease. In another example, a diagnostic nucleotide set that determines continuing graft rejection associated with myocyte necrosis (> grade I) is used to determine that a patient is not receiving adequate treatment under the current treatment regimen. After increased or altered immunosuppressive therapy, diagnostic profiling is conducted to

determine whether continuing graft rejection is progressing. In yet another example, a diagnostic nucleotide set(s) that determine a patient's rejection status and diagnose cytomegalovirus infection is used to balance immunosuppressive and anti-viral therapy.

Example 12: Identification of diagnostic nucleotide sets for kidney and liver allograft rejection

Diagnostic tests for rejection are identified using patient leukocyte expression profiles to identify a molecular signature correlated with rejection of a transplanted kidney or liver. Blood, or other leukocyte source, samples are obtained from patients undergoing kidney or liver biopsy following liver or kidney transplantation, respectively. Such results reveal the histological grade, i.e., the state and severity of allograft rejection. Expression profiles are obtained from the samples as described above, and the expression profile is correlated with biopsy results. In the case of kidney rejection, clinical data is collected corresponding to urine output, level of creatine clearance, and level of serum creatine (and other markers of renal function). Clinical data collected for monitoring liver transplant rejection includes, biochemical characterization of serum markers of liver damage and function such as SGOT, SGPT, Alkaline phosphatase, GGT, Bilirubin, Albumin and Prothrombin time.

Leukocyte nucleotide sequence expression profiles are collected and correlated with important clinical states and outcomes in renal or hepatic transplantation. Examples of useful clinical correlates are given here:

1. Rejection episode of at least moderate histologic grade, which results in treatment of the patient with additional corticosteriods, anti-T cell antibodies, or total lymphoid irradiation.
2. The absence of histologic rejection and normal or unchanged allograft function (based on tests of renal or liver function listed above).
3. The presence of severe allograft dysfunction or worsening allograft dysfunction during the study period (based on tests of renal and hepatic function listed above).
4. Documented CMV infection by culture, histology, or PCR, and at least one clinical sign or symptom of infection.
5. Specific graft biopsy rejection grades
6. Rejection of mild to moderate histologic severity prompting augmentation of the patient's chronic immunosuppressive regimen
7. Infections other than CMV, esp. Epstein Barr virus (EBV)
8. Lymphoproliferative disorder (also called, post-transplant lymphoma)
9. Graft Failure or Retransplantation
10. Need for hemodialysis or other renal replacement therapy for renal transplant patients.

11. Hepatic encephalopathy for liver transplant recipients.
12. All cause mortality

Subsets of the candidate library (or of a previously identified diagnostic nucleotide set), are identified, according to the above procedures, that have predictive and/or diagnostic value for kidney or liver allograft rejection.

Example 13: Identification of diagnostic nucleotide sequences sets for use in the diagnosis, prognosis, risk stratification, and treatment of Atherosclerosis, Stable Angina Pectoris, and acute coronary syndrome.

Prediction of complications of atherosclerosis: angina pectoris.

Over 50 million in the US have atherosclerotic coronary artery disease (CAD). Almost all adults have some atherosclerosis. The most important question is who will develop complications of atherosclerosis. Patients with angiographically-confirmed atherosclerosis are enrolled in a study, and followed over time. Leukocyte expression profiles are taken at the beginning of the study, and routinely thereafter. Some patients develop angina and others do not. Expression profiles are correlated with development of angina, and subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive and/or diagnostic value for angina pectoris.

Alternatively, patients are followed by serial angiography. Profiles are collected at the first angiography, and at a repeat angiography at some future time (for example, after 1 year). Expression profiles are correlated with progression of disease, measured, for example, by decrease in vessel lumen diameter. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive and/or diagnostic value for progression of atherosclerosis.

Prediction and/or diagnosis of acute coronary syndrome

The main cause of death due to coronary atherosclerosis is the occurrence of acute coronary syndromes: myocardial infarction and unstable angina. Patients with at a very high risk of acute coronary syndrome (e.g., patients with a history of acute coronary syndrome, patients with atherosclerosis, patients with multiple traditional risk factors, clotting disorders or lupus) are enrolled in a prospective study. Leukocyte expression profiles are taken at the beginning of the study period and patients are monitored for the occurrence of unstable angina and/or myocardial

infarction. Standard criteria for the occurrence of an event are used (serum enzyme elevation, EKG, nuclear imaging or other), and the occurrence of these events can be collected from the patient, the patient's physician, the medical record or medical database. Expression profiles (taken at the beginning of the study) are correlated with the occurrence of an acute event. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for occurrence of an acute event.

In addition, expression profiles (taken at the time that an acute event occurs) are correlated with the occurrence of an acute event. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have diagnostic value for occurrence of an acute event.

Risk stratification: occurrence of coronary artery disease

The established and classic risks for the occurrence of coronary artery disease and complications of that disease are: cigarette smoking, diabetes, hypertension, hyperlipidemia and a family history of early atherosclerosis. Obesity, sedentary lifestyle, syndrome X, cocaine use, chronic hemodialysis and renal disease, radiation exposure, endothelial dysfunction, elevated plasma homocysteine, elevated plasma lipoprotein a, and elevated CRP. Infection with CMV and chlamydia infection are less well established, controversial or putative risk factors for the disease. These risk factors can be assessed or measured in a population.

Leukocyte expression profiles are measured in a population possessing risk factors for the occurrence of coronary artery disease. Expression profiles are correlated with the presence of one or more risk factors (that may correlate with future development of disease and complications). Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for the development of coronary artery disease.

Additional examples of useful correlation groups in cardiology include:

1.Samples from patients with a high risk factor burden (e.g., smoking, diabetes, high cholesterol, hypertension, family history) versus samples from those same patients at different times with fewer risks, or versus samples from different patients with fewer or different risks.

2.Samples from patients during an episode of unstable angina or myocardial infarction versus paired samples from those same patients before the episode or after recovery, or from different patients without these diagnoses.

3.Samples from patients (with or without documented atherosclerosis) who subsequently develop clinical manifestations of atherosclerosis such as stable angina, unstable angina, myocardial infarction, or stroke ,versus samples from patients (with or without atherosclerosis) who do not develop these manifestations over the same time period.

4.Samples from patients who subsequently respond to a given medication or treatment regimen versus samples from those same or different patients who subsequently do not respond to a given medication or treatment regimen.

Example 14: Identification of diagnostic nucleotide sets for use in diagnosing and treating Restenosis

Restenosis is the re-narrowing of a coronary artery after an angioplasty. Patients are identified who are about to, or have recently undergone angioplasty. Leukocyte expression profiles are measured before the angioplasty, and at 1 day and 1-2 weeks after angioplasty or stent placement. Patients have a follow-up angiogram at 3 months and/or are followed for the occurrence of clinical restenosis, e.g., chest pain due to re-narrowing of the artery, that is confirmed by angiography. Expression profiles are compared between patients with and without restenosis, and candidate nucleotide profiles are correlated with the occurrence of restenosis. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for the development of restenosis.

Example 15: Identification of diagnostic nucleotide sets for use in monitoring treatment and/or progression of Congestive Heart Failure

CHF effects greater than 5 million individuals in the US and the prevalence of this disorder is growing as the population ages. The disease is chronic and debilitating. Medical expenditures are huge due to the costs of drug treatments, echocardiograms and other tests, frequent hospitalization and cardiac transplantation. The primary causes of CHF are coronary artery disease, hypertension and idiopathic

cardiomyopathy. Congestive heart failure is the number one indication for heart transplantation.

There is ample recent evidence that congestive heart failure is associated with systemic inflammation. A leukocyte test with the ability to determine the rate of progression and the adequacy of therapy is of great interest. Patients with severe CHF are identified, e.g. in a CHF clinic, an inpatient service, or a CHF study or registry (such as the cardiac transplant waiting list/registry). Expression profiles are taken at the beginning of the study and patients are followed over time, for example, over the course of one year, with serial assessments performed at least every three months. Further profiles are taken at clinically relevant end-points, for example: hospitalization for CHF, death, pulmonary edema, worsening of Ejection Fraction or increased cardiac chamber dimensions determined by echocardiography or another imaging test, and/or exercise testing of hemodynamic measurements. Clinical data is collected from patients if available, including:

Serial C-Reactive Protein (CRP), other serum markers, echocardiography (e.g., ejection fraction or another echocardiographic measure of cardiac function), nuclear imaging, NYHA functional classes, hospitalizations for CHF, quality of life measures, renal function, transplant listing, pulmonary edema, left ventricular assist device use, medication use and changes.

Expression profiles correlating with progression of CHF are identified. Expression profiles predicting disease progression, monitoring disease progression and response to treatment, and predicting response to a particular treatment(s) or class of treatment(s) are identified. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for the progression of CHF. Such diagnostic nucleotide sets are also useful for monitoring response to treatment for CHF.

Example 16: Identification of diagnostic nucleotide sets for use in monitoring treatment and/or progression of Rheumatoid arthritis

Rheumatoid arthritis (hereinafter, "RA") is a chronic and debilitating inflammatory arthritis. The diagnosis of RA is made by clinical criteria and radiographs. A new class of medication, TNF blockers, are effective, but the drugs are expensive, have side effects and not all patients respond to treatment. In addition, relief of disease symptoms does not always correlate with inhibition of joint

destruction. For these reasons, an alternative mechanism for the titration of therapy is needed.

An observational study was conducted in which a cohort of patients meeting American College of Rheumatology (hereinafter "ARC") criteria for the diagnosis of RA was identified. Arnett et al. (1988) Arthritis Rheum 31:315-24. Patients gave informed consent and a peripheral blood mononuclear cell RNA sample was obtained by the methods as described herein. When available, RNA samples were also obtained from surgical specimens of bone or synovium from effected joints, and synovial fluid .

From each patient, the following clinical information was obtained if available:

Demographic information; information relating to the ACR criteria for RA; presence or absence of additional diagnoses of inflammatory and non-inflammatory conditions; data from laboratory test, including complete blood counts with differentials, CRP, ESR, ANA, Serum IL6, Soluble CD40 ligand, LDL, HDL, Anti-DNA antibodies, rheumatoid factor, C3, C4, serum creatinine and any medication levels; data from surgical procedures such as gross operative findings and pathological evaluation of resected tissues and biopsies; information on pharmacological therapy and treatment changes; clinical diagnoses of disease "flare"; hospitalizations; quantitative joint exams; results from health assessment questionnaires (HAQs); other clinical measures of patient symptoms and disability; physical examination results and radiographic data assessing joint involvement, synovial thickening, bone loss and erosion and joint space narrowing and deformity.

From these data, measures of improvement in RA are derived as exemplified by the ACR 20% and 50% response/improvement rates (Felson et al. 1996). Measures of disease activity over some period of time is derived from these data as are measures of disease progression. Serial radiography of effected joints is used for objective determination of progression (e.g., joint space narrowing, peri-articular osteoporosis, synovial thickening). Disease activity is determined from the clinical scores, medical history, physical exam, lab studies, surgical and pathological findings. The collected clinical data (disease criteria) is used to define patient or sample groups for correlation of expression data. Patient groups are identified for comparison, for example, a patient group that possesses a useful or interesting clinical distinction, verses a patient group that does not possess the distinction. Examples of useful and

interesting patient distinctions that can be made on the basis of collected clinical data are listed here:

1. Samples from patients during a clinically diagnosed RA flare versus samples from these same or different patients while they are asymptomatic.

2. Samples from patients who subsequently have high measures of disease activity versus samples from those same or different patients who have low subsequent disease activity.

3. Samples from patients who subsequently have high measures of disease progression versus samples from those same or different patients who have low subsequent disease progression.

4. Samples from patients who subsequently respond to a given medication or treatment regimen versus samples from those same or different patients who subsequently do not respond to a given medication or treatment regimen (for example, TNF pathway blocking medications).

5. Samples from patients with a diagnosis of osteoarthritis versus patients with rheumatoid arthritis.

6. Samples from patients with tissue biopsy results showing a high degree of inflammation versus samples from patients with lesser degrees of histological evidence of inflammation on biopsy.

Expression profiles correlating with progression of RA are identified. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for the progression of RA.

Diagnostic nucleotide set(s) are identified which predict respond to TNF blockade. Patients are profiled before and during treatment with these medications. Patients are followed for relief of symptoms, side effects and progression of joint destruction, e.g., as measured by hand radiographs. Expression profiles correlating with response to TNF blockade are identified. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures that have predictive value for response to TNF blockade.

Example 17: Identification of diagnostic nucleotide sets for diagnosis of Systemic Lupus Erythematosis

SLE is a chronic, systemic inflammatory disease characterized by dysregulation of the immune system. Clinical manifestations affect every organ system and include skin rash, renal dysfunction, CNS disorders, arthralgias and hematologic abnormalities. SLE clinical manifestations tend to both recur intermittently (or "flare") and progress over time, leading to permanent end-organ damage.

An observational study was conducted in which a cohort of patients meeting American College of Rheumatology (hereinafter "ACR") criteria for the diagnosis of SLE were identified. See Tan et al. (1982) Arthritis Rheum 25:1271-7. Patients gave informed consent and a peripheral blood mononuclear cell RNA sample was obtained by the methods as described herein.

From each patient, the following clinical information was obtained if available:

Demographic information, ACR criteria for SLE, additional diagnoses of inflammatory and non-inflammatory conditions, data from laboratory testing including complete blood counts with differentials, CRP, ESR, ANA, Serum IL6, Soluble CD40 ligand, LDL, HDL, Anti-DNA antibodies, rheumatoid factor, C3, C4, serum creatinine (and other measures of renal dysfunction) and any medication levels, data from surgical procedures such as gross operative findings and pathological evaluation of resected tissues and biopsies (e.g., renal, CNS), information on pharmacological therapy and treatment changes, clinical diagnoses of disease "flare", hospitalizations, quantitative joint exams, results from health assessment questionnaires (HAQs), SLEDAIs (a clinical score for SLE activity that assess many clinical variables), other clinical measures of patient symptoms and disability, physical examination results and carotid ultrasonography.

The collected clinical data (disease criteria) is used to define patient or sample groups for correlation of expression data. Patient groups are identified for comparison, for example, a patient group that possesses a useful or interesting clinical distinction, verses a patient group that does not possess the distinction. Measures of disease activity in SLE are derived from the clinical data described above to divide patients (and patient samples) into groups with higher and lower disease activity over some period of time or at any one point in time. Such data are SLEDAI scores and

other clinical scores, levels of inflammatory markers or complement, number of hospitalizations, medication use and changes, biopsy results and data measuring progression of end-organ damage or end-organ damage, including progressive renal failure, carotid atherosclerosis, and CNS dysfunction. Further examples of useful and interesting patient distinctions that can be made on the basis of collected clinical data are listed here:

Samples from patients during a clinically diagnosed SLE flare versus samples from these same or different patients while they are asymptomatic or while they have a documented infection.

1. Samples from patients who subsequently have high measures of disease activity versus samples from those same or different patients who have low subsequent disease activity.

2. Samples from patients who subsequently have high measures of disease progression versus samples from those same or different patients who have low subsequent disease progression.

3. Samples from patients who subsequently respond to a given medication or treatment regimen versus samples from those same or different patients who subsequently do not respond to a given medication or treatment regimen.

4. Samples from patients with premature carotid atherosclerosis on ultrasonography versus patients with SLE without premature atherosclerosis.

Expression profiles correlating with progression of SLE are identified, including expression profiles corresponding to end-organ damage and progression of end-organ damage. Expression profiles are identified predicting disease progression or disease "flare", response to treatment or likelihood of response to treatment, predict likelihood of "low" or "high" disease measures (optionally described using the SLEDAI score), and presence or likelihood of developing premature carotid atherosclerosis. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for the progression of SLE.

Example 18: Identification of a diagnostic nucleotide set for diagnosis of cytomegalovirus

Cytomegalovirus is a very important cause of disease in immunosupressed patients, for example, transplant patients, cancer patients, and AIDS patients. The virus can cause inflammation and disease in almost any tissue (particularly the colon, lung, bone marrow and retina). It is increasingly important to identify patients with current or impending clinical CMV disease, particularly when immunosuppressive drugs are to be used in a patient, e.g. for preventing transplant rejection.

Leukocytes are profiled in patients with active CMV, impending CMV, or no CMV. Expression profiles correlating with diagnosis of active or impending CMV are identified. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for the diagnosis of active or impending CMV. Diagnostic nucleotide set(s) identified with predictive value for the diagnosis of active or impending CMV may be combined, or used in conjunction with, cardiac, liver and/or kidney allograft-related diagnostic gene set(s) (described in Examples 11 and 12).

In addition, or alternatively, CMV nucleotide sequences are obtained, and a diagnostic nucleotide set is designed using CMV nucleotide sequence. The entire sequence of the organism is known and all CMV nucleotide sequences can be isolated and added to the library using the sequence information and the approach described below. Known expressed genes are preferred. Alternatively, nucleotide sequences are selected to represent groups of CMV genes that are coordinately expressed (immediate early genes, early genes, and late genes) (Spector et al. 1990, Stamminger et al. 1990).

CMV nucleotide sequences were isolated as follows: Primers were designed to amplify known expressed CMV genes, based on the publically available sequence of CMV strain AD 169 (Genbank LOCUS: HEHCMVCG 229354 bp; DEFINITION Human cytomegalovirus strain AD169 complete genome; ACCESSION X17403; VERSION X17403.1 GI:59591). The following primer were used to PCR amplify nucleotide sequences from 175 ng of AD 169 viral genomic DNA (Advance Biotechnologies Incorporated) as a template:

CMV GENE	PRIMER SEQUENCES	SEQ. ID. NO:
UL21 5'	atgtggccgcttctgaaaaac	8771

UL21 3'	tcatggggtggggacgggg	8772
UL33 5'	gtacgcgctgctgggtcatg	8773
UL33 3'	tcataccccgctgaggttatg	8774
UL54 5'	cacggacgacgacgctgacg	8775
UL54 3'	gtacggcagaaaagccggctc	8776
UL55 5'	caccaaagacacgtcgttacag	8777
UL55 3'	tcagacgttctcttcttcgtcg	8778
UL75 5'	cagcggcgctcaacatttcac	8779
UL75 3'	tcagcatgtcttgagcatgcgg	8780
UL80 5'	cctccccaactactactaccg	8781
UL80 3'	ttactcgagcttattgagcgcag	8782
UL83 5'	cacgtcgggcgttatgacac	8783
UL83 3'	tcaacctcggtgctttttggg	8784
UL97 5'	ctgtctgctcattctggcgg	8785
UL97 3'	ttactcggggaacagttggcg	8786
UL106 5'	atgatgaccgaccgcacgga	8787
UL106 3'	tcacggtggctcgatacactg	8788
UL107 5'	aagcttccttacagcataactgt	8789
UL107 3'	ccttataacatgtattttgaaaaattg	8790
UL109 5'	atgatacacgactaccactgg	8791
UL109 3'	ttacgagcaagagttcatcacg	8792
UL112 5'	ctgcgtgtcctcgctgggt	8793
UL112 3'	tcacgagtccactcggaaagc	8794
UL113 5'	ctcgtcttcttcggctccac	8795
UL113 3'	ttaatcgtcgaaaaacgccgcg	8796
UL122 5'	gatgcttgtaacgaaggcgtc	8797
UL122 3'	ttactgagacttgttcctcagg	8798
UL123 5'	gtagcctacactttggccacc	8799
UL123 3'	ttactggtcagccttgcttcta	8800
IRL2 5'	acgtccctggtagacggg	8801
IRL2 3'	ttataagaaaagaagcacaagctc	8802
IRL3 5'	atgtattgttttctttttttacagaaag	8803
IRL3 3'	ttatattattatcaaaacgaaaaacag	8804
IRL4 5'	cttctcctttccttaatctcgg	8805
IRL4 3'	ctatacggagatcgcggtcc	8806
IRL5 5'	atgcatacatacacgcgtgcat	8807
IRL5 3'	ctaccatataaaaacgcagggg	8808
IRL7 5'	atgaaagcaagaggcagccg	8809
IRL7 3'	tcataaggtaacgatgctacttt	8810
IRL13 5'	atggactggcgatttacggtt	8811
IRL13 3'	ctacattgtgccatttctcagt	8812
US2 5'	atgaacaatctctggaaagcctg	8813
US2 3'	tcagcacacgaaaaaccgcatc	8814

US3 5'	atgaagccggtgttggtgctc	8815
US3 3'	ttaaataaatcgcagacgggcg	8816
US6 5'	atggatctcttgattcgtctcg	8817
US6 3'	tcaggagccacaacgtcgaatc	8818
US11 5'	cgcaaaacgctactggctcc	8819
US11 3'	tcaccactggtccgaaaacatc	8820
US18 5'	tacggctggtccgtcatcgt	8821
US18 3'	ttacaacaagctgaggagactc	8822
US27 5'	atgaccacctctacaaataatcaaac	8823
US27 3'	gtagaaacaagcgttgagtccc	8824
US28 5'	cgttgcggtgtctcagtcg	8825
US28 3'	tcatgctgtggtaccaggata	8826

The PCR reaction conditions were 10 mM Tris pH 8.3, 3.5 mM MgCl2, 25 mM KCl, 200 uM dNTP's, 0.2 uM primers, and 5 Units of Taq Gold. The cycle parameters were as follows:

1. 95°C for 30 sec
2. 95°C for 15 sec
3. 56°C for 30 sec
4. 72°C for 2 min
5. go to step 2, 29 times
6. 72°C for 2 min
7. 4°C forever

PCR products were gel purified, and DNA was extracted from the agarose using the QiaexII gel purification kit (Qiagen). PCR product was ligated into the T/A cloning vector p-GEM-T-Easy (Promega) using 3 ul of gel purified PCR product and following the Promega protocol. The products of the ligation reaction were transformed and plated as described in the p-GEM protocol. White colonies were picked and grow culture in LB-AMP medium. Plasmid was prepared from these cultures using Qiagen Miniprep kit (Qiagen). Restriction enzyme digested plasmid (Not I and EcoRI) was examined after agarose gel electrophoresis to assess insert size. When the insert was the predicted size, the plasmid was sequenced by well-known techniques to confirm the identity of the CMV gene. Using forward and reverse primers that are complimentary to sequences flanking the insert cloning site (M13F and M13R), the isolated CMV gene was amplified and purified as described above.

Amplified cDNAs were used to create a microarray as described above. In addition, 50mer oligonucleotides corresponding the CMV genes listed above were designed, synthesized and placed on a microarray using methods described elsewhere in the specification.

Alternatively, oligonucleotide sequences aredesigned and synthesized for oligonucleotide array expression analysis from CMV genes as described in examples 20-22.

Diagnostic nucleotide set(s) for expression of CMV genes is used in combination with diagnostic leukocyte nucleotide sets for diagnosis of other conditions, e.g. organ allograft rejection.

Example 19: Identification of diagnostic nucleotide sets for monitoring response to Statins

HMG-CoA reductase inhibitors, called "Statins," are very effective in preventing complications of coronary artery disease in either patients with coronary disease and high cholesterol (secondary prevention) or patients without known coronary disease and with high cholesterol (primary prevention). Examples of Statins are (generic names given) pravistatin, atorvastatin, and simvastain. Monitoring response to Statin therapy is of interest. Patients are identified who are on or are about to start Statin therapy. Leukocytes are profiled in patients before and after initiation of therapy, or in patients already being treated with Statins. Data is collected corresponding to cholesterol level, markers of inflammation (e.g., C-Reactive Protein and the Erythrocyte Sedimentation Rate), measures of endothelial function (e.g., improved forearm resistance or coronary flow reserve) and clinical endpoints (new stable angina, unstable angina, myocardial infarction, ventricular arrhythmia, claudication). Patient groups can be defined based on their response to Statin therapy (cholesterol, clinical endpoints, endothelial function). Expression profiles correlating with response to Statin treatment are identified. Subsets of the candidate library (or a previously identified diagnostic nucleotide set) are identified, according to the above procedures, that have predictive value for the response to Statins. Members of candidate nucleotide sets with expression that is altered by Statins are disease target nucleotides sequences.

Example 20--Probe Selection for a 24,000 Feature Array

This Example describes the compilation of almost 8,000 unique genes and ESTs using sequences identified from the sources described below. The sequences of these genes and ESTs were used to design probes, as described in the following Example.

Tables 3A, 3B and 3C list the sequences identified in the subtracted leukocyte expression libraries. All sequences that were identified as corresponding to a known RNA transcript were represented at least once, and all unidentified sequences were represented twice – once by the sequence on file and again by the complementary sequence – to ensure that the sense (or coding) strand of the gene sequence was included.

Table 3A. Table 3A contained all those sequences in BioCardia's subtracted libraries that matched sequences in GenBank's nr, EST_Human, and UniGene databases with an acceptable level of confidence. All the entries in the table representing the sense strand of their genes were grouped together and all those representing the antisense strand were grouped. A third group contained those entries whose strand could not be determined. Two complementary probes were designed for each member of this third group.

Table 3B and 3C. Table 3B and 3C contained all those sequences in the leukocyte expression subtracted library that did not match sequences in GenBank's nr, EST_Human, and UniGene databases with an acceptable level of confidence, but which had a high probability of representing real mRNA sequences. Sequences in Table 3B did not match anything in the databases above but matched regions of the human genome draft and were spatially clustered along it, suggesting that they were exons, rather than genomic DNA included in the library by chance. Sequences in Table 3C also aligned well to regions of the human genome draft, but the aligned regions were interrupted by genomic DNA, meaning they were likely to be spliced transcripts of multiple exon genes.

Table 3B lists 510 clones and Table 3C lists 48 clones that originally had no similarity with any sequence in the public databases. Blastn searches conducted after the initial filing have identified sequences in the public database with high similarity (E values less than 1e-40) to the sequences determined for these clones. Table 3B contained 272 clones and Table 3C contained 25 clones that were found to have high similarity to sequences in dbEST. The sequences of the similar dbEST clones were

used to design probes. Sequences from clones that contained no similar regions to any sequence in the database were used to design a pair of complementary probes.

Probes were designed from database sequences that had the highest similarity to each of the sequenced clones in Tables 3A, 3B, and 3C. Based on BLASTn searches the most similar database sequence was identified by locus number and the locus number was submitted to GenBank using batch Entrez (http://www.ncbi.nlm.nih.gov/entrez/batchentrez.cgi?db=Nucleotide) to obtain the sequence for that locus. The GenBank entry sequence was used because in most cases it was more complete or was derived from multi-pass sequencing and thus would likely have fewer errors than the single pass cDNA library sequences. When only UniGene cluster IDs were available for genes of interest, the respective sequences were extracted from the UniGene_unique database, build 137, downloaded from NCBI (ftp://ncbi.nlm.nih.gov/repository/UniGene/). This database contains one representative sequence for each cluster in UniGene.

Summary of BioCardia library clones used in probe design.

Table	Sense Strand	Antisense Strand	Strand Undetermined
Table 3A	3621	763	124
Table 3B	142	130	238
Table 3C	19	6	23
Totals	**3782**	**899**	**385**

Literature Searches

Example 2 describes searches of literature databases. We also searched for research articles discussing genes expressed only in leukocytes or involved in inflammation and particular disease conditions, including genes that were specifically expressed or down-regulated in a disease state. Searches included, but were not limited to, the following terms and various combinations of theses terms: inflammation, atherosclerosis, rheumatoid arthritis, osteoarthritis, lupus, SLE, allograft, transplant, rejection, leukocyte, monocyte, lymphocyte, mononuclear, macrophage, neutrophil, eosinophil, basophil, platelet, congestive heart failure, expression, profiling, microarray, inflammatory bowel disease, asthma, RNA expression, gene expression, granulocyte.

A UniGene cluster ID or GenBank accession number was found for each gene in the list. The strand of the corresponding sequence was determined, if possible, and the genes were divided into the three groups: sense (coding) strand, anti-sense strand, or strand unknown. The rest of the probe design process was carried out as described above for the sequences from the leukocyte subtracted expression library.

Database Mining

Database mining was performed as described in Example 2. In addition, the Library Browser at the NCBI UniGene web site (http://www.ncbi.nlm.nih.gov/UniGene/lbrowse.cgi?ORG=Hs&DISPLAY=ALL) was used to identify genes that are specifically expressed in leukocyte cell populations. All expression libraries available at the time were examined and those derived from leukocytes were viewed individually. Each library viewed through the Library Browser at the UniGene web site contains a section titled "Shown below are UniGene clusters of special interest only" that lists genes that are either highly represented or found only in that library. Only the genes in this section were downloaded from each library. Alternatively, every sequence in each library is downloaded and then redundancy between libraries is reduced by discarding all UniGene cluster IDs that are represented more than once.

A total of 439 libraries were downloaded, containing 35,819 genes, although many were found in more than one library. The most important libraries from the remaining set were separated and 3,914 genes remained. After eliminating all redundancy between these libraries and comparing the remaining genes to those listed in Tables 3A, 3B and 3C, the set was reduced to 2,573 genes in 35 libraries (listed below). From these, all genes in first 30 libraries were used to design probes. A random subset of genes was used from Library Lib.376, "Activated_T-cells_XX". From the last four libraries, a random subset of sequences listed as "ESTs, found only in this library" was used.

Library ID	Library Name	Category	No. of sequences before reduction	No. of sequences used on array*
Lib.2228	Human_leukocyte_MATCHMAKER_cDNA_Library	other/unclassified	4	3

Lib.238	RA-MO-III (activated monocytes from RA patient)	Blood	2	1
Lib.242	Human_peripheral_blood_(Whole)_(Steve_Elledge)	Blood	4	2
Lib.2439	Subtracted_cDNA_libraries_from_human_Jurkat_cells	other/unclassified	4	1
Lib.323	Activated_T-cells_I	other/unclassified	19	3
Lib.327	Monocytes,_stimulated_II	Blood	92	35
Lib.387	Macrophage_I	other/unclassified	84	24
Lib.409	Activated_T-cells_IV	other/unclassified	37	10
Lib.410	Activated_T-cells_VIII	other/unclassified	27	10
Lib.411	Activated_T-cells_V	other/unclassified	41	9
Lib.412	Activated_T-cells_XII	other/unclassified	29	12
Lib.413	Activated_T-cells_XI	other/unclassified	13	6
Lib.414	Activated_T-cells_II	other/unclassified	69	30
Lib.429	Macrophage_II	other/unclassified	56	24
Lib.4480	Homo_sapiens_rheumatoid_arthritis_fibroblast-like_synovial	other/unclassified	7	6
Lib.476	Macrophage,_subtracted_(total_cDNA)	other/unclassified	11	1
Lib.490	Activated_T-cells_III	other/unclassified	9	5
Lib.491	Activated_T-cells_VII	other/unclassified	27	8
Lib.492	Activated_T-cells_IX	other/unclassified	16	5
Lib.493	Activated_T-cells_VI	other/unclassified	31	15
Lib.494	Activated_T-cells_X	other/unclassified	18	5
Lib.498	RA-MO-I (activated peripheral blood monocytes from RA patient)	Blood	2	1
Lib.5009	Homo_Sapiens_cDNA_Library_from_Peripheral_White_Blood_Cell	other/unclassified	3	3
Lib.6338	human_activated_B_lymphocyte	Tonsils	9	8
Lib.6342	Human_lymphocytes	other/unclassified	2	2
Lib.646	Human_leukocyte_(M.L.Markelov)	other/unclassified	1	1
Lib.689	Subtracted_cDNA_library_of_activated_B_lymphocyte	Tonsil	1	1
Lib.773	PMA-induced_HL60_cell_subtraction_library (leukemia)	other/unclassified	6	3
Lib.1367	cDNA_Library_from_rIL-2_activated_lymphocytes	other/unclassified	3	2
Lib.5018	Homo_sapiens_CD4+_T-cell_clone_HA1.7	other/unclassified	6	3
Lib.376	Activated_T-cells_XX	other/unclassified	999	119
Lib.669	NCI_CGAP_CLL1 (Lymphocyte)	Blood	353	81†
Lib.1395	NCI_CGAP_Sub6 (germinal center b-cells)	B cells germinal	389	100†
Lib.2217	NCI_CGAP_Sub7 (germinal center b-cells)	B cells germinal	605	200†
Lib.289	NCI_CGAP_GCB1 (germinal center b-cells)	Tonsil	935	200†
Total			**3,914**	**939**

* Redundancy of UniGene numbers between the libraries was eliminated.

† A subset of genes flagged as "Found only in this library" were taken.

Angiogenesis Markers

215 sequences derived from an angiogenic endothelial cell subtracted cDNA library obtained from Stanford University were used for probe design. Briefly, using well known subtractive hybridization procedures, (as described in, e.g., US Patent Numbers 5,958,738; 5,589,339; 5,827,658; 5,712,127; 5,643,761; 5,565,340) modified to normalize expression by suppressing over-representation of abundant RNA species while increasing representation of rare RNA species, a library was produced that is enriched for RNA species (messages) that are differentially expressed between test (stimulated) and control (resting) HUVEC populations. The subtraction/suppression protocol was performed as described by the kit manufacturer (Clontech, PCR-select cDNA Subtraction Kit).

Pooled primary HUVECs (Clonetics) were cultured in 15% FCS, M199 (GibcoBRL) with standard concentrations of Heparin, Penicillin, Streptomycin, Glutamine and Endothelial Cell Growth Supplement. The cells were cultured on 1% gelatin coated 10 cm dishes. Confluent HUVECs were photographed under phase contrast microscopy. The cells formed a monolayer of flat cells without gaps. Passage 2-5 cells were used for all experiments. Confluent HUVECs were treated with trypsin/EDTA and seeded onto collagen gels. Collagen gels were made according to the protocol of the Collagen manufacturer (Becton Dickinson Labware). Collagen gels were prepared with the following ingredients: Rat tail collagen type I (Collaborative Biomedical) 1.5 mg/mL, mouse laminin (Collaborative Biomedical) 0.5 mg/mL, 10% 10X media 199 (Gibco BRL). 1N NaOH, 10 X PBS and sterile water were added in amounts recommended in the protocol. Cell density was measured by microscopy. 1.2×10^6 cells were seeded onto gels in 6-well, 35 mm dishes, in 5% FCS M199 media. The cells were incubated for 2 hrs at 37 C with 5% CO2. The media was then changed to the same media with the addition of VEGF (Sigma) at 30ng/mL media. Cells were cultured for 36 hrs. At 12, 24 and 36 hrs, the cells were observed with phase contrast microscopy. At 36 hours, the cells were observed elongating, adhering to each other and forming lumen structures. At 12 and 24 hrs media was aspirated and refreshed. At 36 hrs, the media was aspirated, the cells were rinsed with PBS and then treated with Collagenase (Sigma) 2.5mg/mL PBS for 5 min with active agitation until the collagen gels were liquefied. The cells were then centrifuged at 4C, 2000g for 10 min. The supernatant was removed and the cells

were lysed with 1 mL Trizol Reagent (Gibco) per $5x10^6$ cells. Total RNA was prepared as specified in the Trizol instructions for use. mRNA was then isolated as described in the micro-fast track mRNA isolation protocol from Invitrogen. This RNA was used as the tester RNA for the subtraction procedure.

Ten plates of resting, confluent, p4 HUVECs, were cultured with 15 % FCS in the M199 media described above. The media was aspirated and the cells were lysed with 1 mL Trizol and total RNA was prepared according to the Trizol protocol. mRNA was then isolated according to the micro-fast track mRNA isolation protocol from Invitrogen. This RNA served as the control RNA for the subtraction procedure.

The entire subtraction cloning procedure was carried out as per the user manual for the Clontech PCR Select Subtraction Kit. The cDNAs prepared from the test population of HUVECs were divided into "tester" pools, while cDNAs prepared from the control population of HUVECs were designated the "driver" pool. cDNA was synthesized from the tester and control RNA samples described above. Resulting cDNAs were digested with the restriction enzyme RsaI. Unique double-stranded adapters were ligated to the tester cDNA. An initial hybridization was performed consisting of the tester pools of cDNA (with its corresponding adapter) and an excess of the driver cDNA. The initial hybridization results in a partial normalization of the cDNAs such that high and low abundance messages become more equally represented following hybridization due to a failure of driver/tester hybrids to amplify.

A second hybridization involved pooling unhybridized sequences from the first hybridization together with the addition of supplemental driver cDNA. In this step, the expressed sequences enriched in the two tester pools following the initial hybridization can hybridize. Hybrids resulting from the hybridization between members of each of the two tester pools are then recovered by amplification in a polymerase chain reaction (PCR) using primers specific for the unique adapters. Again, sequences originating in a tester pool that form hybrids with components of the driver pool are not amplified. Hybrids resulting between members of the same tester pool are eliminated by the formation of "panhandles" between their common 5' and 3' ends. This process is illustrated schematically in Figure 3. The subtraction was done in both directions, producing two libraries, one with clones that are upregulated in tube-formation and one with clones that are down-regulated in the process.

The resulting PCR products representing partial cDNAs of differentially expressed genes were then cloned (i.e., ligated) into an appropriate vector according to the manufacturer's protocol (pGEM-Teasy from Promega) and transformed into competent bacteria for selection and screening. Colonies (2180) were picked and cultured in LB broth with 50ug/mL ampicillin at 37C overnight. Stocks of saturated LB + 50 ug/mL ampicillin and 15% glycerol in 96-well plates were stored at –80C. Plasmid was prepared from 1.4mL saturated LB broth containing 50 ug/mL ampicillin. This was done in a 96 well format using commercially available kits according to the manufacturer's recommendations (Qiagen 96-turbo prep).

2 probes to represent 22 of these sequences required, therefore, a total of 237 probes were derived from this library.

Viral genes.

Several viruses may play a role in a host of disease including inflammatory disorders, atherosclerosis, and transplant rejection. The table below lists the viral genes represented by oligonucleotide probes on the microarray. Low-complexity regions in the sequences were masked using RepeatMasker before using them to design probes.

Virus	Gene Name	Genome Location
Adenovirus, type 2 Accession #J01917	E1a	1226..1542
	E1b_1	3270...3503
	E2a_2	complement(24089..25885)
	E3-1	27609..29792
	E4 (last exon at 3'-end)	complement(33193..32802)
	IX	3576..4034
	Iva2	complement(4081..5417)
	DNA Polymerase	complement(5187..5418)
Cytomegalovirus (CMV) Accession #X17403	HCMVTRL2 (IRL2)	1893..2240
	HCMVTRL7 (IRL7)	complement(6595..6843)
	HCMVUL21	complement(26497..27024)
	HCMVUL27	complement(32831..34657)
	HCMVUL33	43251..44423
	HCMVUL54	complement(76903..80631)
	HCMVUL75	complement(107901..110132)
	HCMVUL83	complement(119352..121037)
	HCMVUL106	complement(154947..155324)
	HCMVUL109	complement(157514..157810)
	HCMVUL113	161503..162800
	HCMVUL122	complement(169364..170599)
	HCMVUL123 (last exon at 3'-end)	complement(171006..172225)
	HCMVUS28	219200..220171
Epstein-Barr virus (EBV) Accession # NC_001345	Exon in EBNA-1 RNA	67477..67649
	Exon in EBNA-1 RNA	98364..98730
	BRLF1	complement(103366..105183)
	BZLF1 (first of 3 exons)	complement(102655..103155)
	BMLF1	complement(82743..84059)
	BALF2	complement(161384..164770)
Human Herpesvirus 6 (HHV6) Accession #NC_001664	U16/U17	complement(26259..27349)
	U89	complement(133091..135610)
	U90	complement(135664..135948)
	U86	complement(125989..128136)
	U83	123528..123821
	U22	complement(33739..34347)
	DR2 (DR2L)	791..2653
	DR7 (DR7L)	5629..6720
	U95	142941..146306
	U94	complement(141394..142866)
	U39	complement(59588..62080)
	U42	complement(69054..70598)
	U81	complement(121810..122577)

Strand Selection

It was necessary to design sense oligonucleotide probes because the labeling and hybridization protocol to be used with the microarray results in fluorescently-labeled antisense cRNA. All of the sequences we selected to design probes could be divided into three categories:

(1) Sequences known to represent the sense strand

(2) Sequences known to represent the antisense strand

(3) Sequences whose strand could not be easily determined from their descriptions

It was not known whether the sequences from the leukocyte subtracted expression library were from the sense or antisense strand. GenBank sequences are reported with sequence given 5' to 3', and the majority of the sequences we used to design probes came from accession numbers with descriptions that made it clear whether they represented sense or antisense sequence. For example, all sequences containing "mRNA" in their descriptions were understood to be the sequences of the sense mRNA, unless otherwise noted in the description, and all IMAGE Consortium clones are directionally cloned and so the direction (or sense) of the reported sequence can be determined from the annotation in the GenBank record.

For accession numbers representing the sense strand, the sequence was downloaded and masked and a probe was designed directly from the sequence. These probes were selected as close to the 3' end as possible. For accession numbers representing the antisense strand, the sequence was downloaded and masked, and a probe was designed complementary to this sequence. These probes were designed as close to the 5' end as possible (i.e., complementary to the 3' end of the sense strand).

Minimizing Probe Redundancy.

Multiple copies of certain genes or segments of genes were included in the sequences from each category described above, either by accident or by design. Reducing redundancy within each of the gene sets was necessary to maximize the number of unique genes and ESTs that could be represented on the microarray.

Three methods were used to reduce redundancy of genes, depending on what information was available. First, in gene sets with multiple occurrences of one or more

UniGene numbers, only one occurrence of each UniGene number was kept. Next, each gene set was searched by GenBank accession numbers and only one occurrence of each accession number was conserved. Finally, the gene name, description, or gene symbol were searched for redundant genes with no UniGene number or different accession numbers. In reducing the redundancy of the gene sets, every effort was made to conserve the most information about each gene.

We note, however, that the UniGene system for clustering submissions to GenBank is frequently updated and UniGene cluster IDs can change. Two or more clusters may be combined under a new cluster ID or a cluster may be split into several new clusters and the original cluster ID retired. Since the lists of genes in each of the gene sets discussed were assembled at different times, the same sequence may appear in several different sets with a different UniGene ID in each.

Sequences from Table 3A were treated differently. In some cases, two or more of the leukocyte subtracted expression library sequences aligned to different regions of the same GenBank entry, indicating that these sequences were likely to be from different exons in the same gene transcript. In these cases, one representative library sequence corresponding to each presumptive exon was individually listed in Table 3A.

Compilation.

After redundancy within a gene set was sufficiently reduced, a table of approximately 8,000 unique genes and ESTs was compiled in the following manner. All of the entries in Table 3A were transferred to the new table. The list of genes produced by literature and database searches was added, eliminating any genes already contained in Table 3A. Next, each of the remaining sets of genes was compared to the table and any genes already contained in the table were deleted from the gene sets before appending them to the table.

	Probes
BioCardia Subtracted Leukocyte Expression Library	
Table 3A	4,872
Table 3B	796
Table 3C	85
Literature Search Results	494

Database Mining	1,607
Viral genes	
a. CMV	14
b. EBV	6
c. HHV 6	14
d. Adenovirus	8
Angiogenesis markers: 215, 22 of which needed two probes	237
Arabidopsis thaliana genes	10
Total sequences used to design probes	8,143

Example 21- Design of oligonucleotide probes

This section describes the design of four oligonucleotide probes using Array Designer Ver 1.1 (Premier Biosoft International, Palo Alto, CA).

Clone 40H12

Clone 40H12 was sequenced and compared to the nr, dbEST, and UniGene databases at NCBI using the BLAST search tool. The sequence matched accession number NM_002310, a 'curated RefSeq project' sequence, see Pruitt et al. (2000) Trends Genet. 16:44-47, encoding leukemia inhibitory factor receptor (LIFR) mRNA with a reported E value of zero. An E value of zero indicates there is, for all practical purposes, no chance that the similarity was random based on the length of the sequence and the composition and size of the database. This sequence, cataloged by accession number NM_002310, is much longer than the sequence of clone 40H12 and has a poly-A tail. This indicated that the sequence cataloged by accession number NM_002310 is the sense strand and a more complete representation of the mRNA than the sequence of clone 40H12, especially at the 3' end. Accession number "NM_002310" was included in a text file of accession numbers representing sense strand mRNAs, and sequences for the sense strand mRNAs were obtained by uploading a text file containing desired accession numbers as an Entrez search query using the Batch Entrez web interface and saving the results locally as a FASTA file. The following sequence was obtained, and the region of alignment of clone 40H12 is outlined:

CTCTCTCCCAGAACGTGTCTCTGCTGCAAGGCACCGGGCCCTTTCGCTCTGCAGAACTGC
ACTTGCAAGACCATTATCAACTCCTAATCCCAGCTCAGAAAGGGAGCCTCTGCGACTCAT
TCATCGCCCTCCAGGACTGACTGCATTGCACAGATGATGGATATTTACGTATGTTTGAAA
CGACCATCCTGGATGGTGGACAATAAAAGAATGAGGACTGCTTCAAATTTCCAGTGGCTG
TTATCAACATTTATTCTTCTATATCTAATGAATCAAGTAAATAGCCAGAAAAAGGGGGCT
CCTCATGATTTGAAGTGTGTAACTAACAATTTGCAAGTGTGGAACTGTTCTTGGAAAGCA
CCCTCTGGAACAGGCCGTGGTACTGATTATGAAGTTTGCATTGAAAACAGGTCCCGTTCT
TGTTATCAGTTGGAGAAAACCAGTATTAAAATTCCAGCTCTTTCACATGGTGATTATGAA
ATAACAATAAATTCTCTACATGATTTTGGAAGTTCTACAAGTAAATTCACACTAAATGAA
CAAAACGTTTCCTTAATTCCAGATACTCCAGAGATCTTGAATTTGTCTGCTGATTTCTCA
ACCTCTACATTATACCTAAAGTGGAACGACAGGGGTTCAGTTTTTCCACACCGCTCAAAT
GTTATCTGGGAAATTAAAGTTCTACGTAAAGAGAGTATGGAGCTCGTAAAATTAGTGACC
CACAACACAACTCTGAATGGCAAAGATACACTTCATCACTGGAGTTGGGCCTCAGATATG
CCCTTGGAATGTGCCATTCATTTTGTGGAAATTAGATGCTACATTGACAATCTTCATTTT
TCTGGTCTCGAAGAGTGGAGTGACTGGAGCCCTGTGAAGAACATTTCTTGGATACCTGAT
TCTCAGACTAAGGTTTTTCCTCAAGATAAAGTGATACTTGTAGGCTCAGACATAACATTT
TGTTGTGTGAGTCAAGAAAAAGTGTTATCAGCACTGATTGGCCATACAAACTGCCCCTTG
ATCCATCTTGATGGGGAAAATGTTGCAATCAAGATTCGTAATATTTCTGTTTCTGCAAGT
AGTGGAACAAATGTAGTTTTTACAACCGAAGATAACATATTTGGAACCGTTATTTTTGCT
GGATATCCACCAGATACTCCTCAACAACTGAATTGTGAGACACATGATTTAAAAGAAATT
ATATGTAGTTGGAATCCAGGAAGGGTGACAGCGTTGGTGGGCCCACGTGCTACAAGCTAC
ACTTTAGTTGAAAGTTTTTCAGGAAAATATGTTAGACTTAAAAGAGCTGAAGCACCTACA
AACGAAAGCTATCAATTATTATTTCAAATGCTTCCAAATCAAGAAATATATAATTTTACT
TTGAATGCTCACAATCCGCTGGGTCGATCACAATCAACAATTTTAGTTAATATAACTGAA
AAAGTTTATCCCCATACTCCTACTTCATTCAAAGTGAAGGATATTAATTCAACAGCTGTT
AAACTTTCTTGGCATTTACCAGGCAACTTTGCAAAGATTAATTTTTTATGTGAAATTGAA
ATTAAGAAATCTAATTCAGTACAAGAGCAGCGGAATGTCACAATCAAAGGAGTAGAAAAT
TCAAGTTATCTTGTTGCTCTGGACAAGTTAAATCCATACACTCTATATACTTTTCGGATT
CGTTGTTCTACTGAAACTTTCTGGAAATGGAGCAAATGGAGCAATAAAAAACAACATTTA
ACAACAGAAGCCAGTCCTTCAAAGGGGCCTGATACTTGGAGAGAGTGGAGTTCTGATGGA
AAAAATTTAATAATCTATTGGAAGCCTTTACCCATTAATGAAGCTAATGGAAAAATACTT

TCCTACAATGTATCGTGTTCATCAGATGAGGAAACACAGTCCCTTTCTGAAATCCCTGAT
CCTCAGCACAAAGCAGAGATACGACTTGATAAGAATGACTACATCATCAGCGTAGTGGCT
AAAAATTCTGTGGGCTCATCACCACCTTCCAAAATAGCGAGTATGGAAATTCCAAATGAT
GATCTCAAAATAGAACAAGTTGTTGGGATGGGAAAGGGGATTCTCCTCACCTGGCATTAC
GACCCCAACATGACTTGCGACTACGTCATTAAGTGGTGTAACTCGTCTCGGTCGGAACCA
TGCCTTATGGACTGGAGAAAAGTTCCCTCAAACAGCACTGAAACTGTAATAGAATCTGAT
GAGTTTCGACCAGGTATAAGATATAATTTTTTCCTGTATGGATGCAGAAATCAAGGATAT
CAATTATTACGCTCCATGATTGGATATATAGAAGAATTGGCTCCCATTGTTGCACCAAAT
TTTACTGTTGAGGATACTTCTGCAGATTCGATATTAGTAAAATGGGAAGACATTCCTGTG
GAAGAACTTAGAGGCTTTTTAAGAGGATATTTGTTTTACTTTGGAAAAGGAGAAAGAGAC
ACATCTAAGATGAGGGTTTTAGAATCAGGTCGTTCTGACATAAAAGTTAAGAATATTACT
GACATATCCCAGAAGACACTGAGAATTGCTGATCTTCAAGGTAAAACAAGTTACCACCTG
GTCTTGCGAGCCTATACAGATGGTGGAGTGGGCCCGGAGAAGAGTATGTATGTGGTGACA
AAGGAAAATTCTGTGGGATTAATTATTGCCATTCTCATCCCAGTGGCAGTGGCTGTCATT
GTTGGAGTGGTGACAAGTATCCTTTGCTATCGGAAACGAGAATGGATTAAAGAAACCTTC
TACCCTGATATTCCAAATCCAGAAAACTGTAAAGCATTACAGTTTCAAAAGAGTGTCTGT
GAGGGAAGCAGTGCTCTTAAAACATTGGAAATGAATCCTTGTACCCCAAATAATGTTGAG
GTTCTGGAAACTCGATCAGCATTTCCTAAAATAGAAGATACAGAAATAATTTCCCCAGTA
GCTGAGCGTCCTGAAGATCGCTCTGATGCAGAGCCTGAAAACCATGTGGTTGTGTCCTAT
TGTCCACCCATCATTGAGGAAGAAATACCAAACCCAGCCGCAGATGAAGCTGGAGGGACT
GCACAGGTTATTTACATTGATGTTCAGTCGATGTATCAGCCTCAAGCAAAACCAGAAGAA
GAACAAGAAAATGACCCTGTAGGAGGGGCAGGCTATAAGCCACAGATGCACCTCCCCATT
AATTCTACTGTGGAAGATATAGCTGCAGAAGAGGACTTAGATAAAACTGCGGGTTACAGA
CCTCAGGCCAATGTAAATACATGGAATTTAGTGTCTCCAGACTCTCCTAGATCCATAGAC
AGCAACAGTGAGATTGTCTCATTTGGAAGTCCATGCTCCATTAATTCCCGACAATTTTTG
ATTCCTCCTAAAGATGAAGACTCTCCTAAATCTAATGGAGGAGGGTGGTCCTTTACAAAC
TTTTTTCAGAACAAACCAAACGATTAACAGTGTCACCGTGTCACTTCAGTCAGCCATCTC
AATAAGCTCTTACTGCTAGTGTTGCTACATCAGCACTGGGCATTCTTGGAGGGATCCTGT
GAAGTATTGTTAGGAGGTGAACTTCACTACATGTTAAGTTACACTGAAAGTTCATGTGCT
TTTAATGTAGTCTAAAAGCCAAAGTATAGTGACTCAGAATCCTCAATCCACAAAACTCAA
GATTGGGAGCTCTTTGTGATCAAGCCAAAGAATTCTCATGTACTCTACCTTCAAGAAGCA
TTTCAAGGCTAATACCTACTTGTACGTACATGTAAAACAAATCCCGCCGCAACTGTTTTC

TGTTCTGTTGTTTGTGGTTTTCTCATATGTATACTTGGTGGAATTGTAAGTGGATTTGCA
GGCCAGGGAGAAAATGTCCAAGTAACAGGTGAAGTTTATTTGCCTGACGTTTACTCCTTT
CTAGATGAAAACCAAGCACAGATTTTAAAACTTCTAAGATTATTCTCCTCTATCCACAGC
ATTCACAAAAATTAATATAATTTTTAATGTAGTGACAGCGATTTAGTGTTTTGTTTGATA
AAGTATGCTTATTTCTGTGCCTACTGTATAATGGTTATCAAACAGTTGTCTCAGGGGTAC
AAACTTTGAAAACAAGTGTGACACTGACCAGCCCAAATCATAATCATGTTTTCTTGCTGT
GATAGGTTTTGCTTGCCTTTTCATTATTTTTTAGCTTTTATGCTTGCTTCCATTATTTCA
GTTGGTTGCCCTAATATTTAAAATTTACACTTCTAAGACTAGAGACCCACATTTTTTAAA
AATCATTTTATTTTGTGATACAGTGACAGCTTTATATGAGCAAATTCAATATTATTCATA
AGCATGTAATTCCAGTGACTTACTATGTGAGATGACTACTAAGCAATATCTAGCAGCGTT
AGTTCCATATAGTTCTGATTGGATTTCGTTCCTCCTGAGGAGACCATGCCGTTGAGCTTG
GCTACCCAGGCAGTGGTGATCTTTGACACCTTCTGGTGGATGTTCCTCCCACTCATGAGT
CTTTTCATCATGCCACATTATCTGATCCAGTCCTCACATTTTTAAATATAAAACTAAAGA
GAGAATGCTTCTTACAGGAACAGTTACCCAAGGGCTGTTTCTTAGTAACTGTCATAAACT
GATCTGGATCCATGGGCATACCTGTGTTCGAGGTGCAGCAATTGCTTGGTGAGCTGTGCA
GAATTGATTGCCTTCAGCACAGCATCCTCTGCCCACCCTTGTTTCTCATAAGCGATGTCT
GGAGTGATTGTGGTTCTTGGAAAAGCAGAAGGAAAAACTAAAAAGTGTATCTTGTATTTT
CCCTGCCCTCAGGTTGCCTATGTATTTTACCTTTTCATATTTAAGGCAAAAGTACTTGAA
AATTTTAAGTGTCCGAATAAGATATGTCTTTTTTTGTTTGTTTTTTTTTGGTTGGTTGTTTG
TTTTTTATCATCTGAGATTCTGTAATGTATTTGCAAATAATGGATCAATTAATTTTTTTT
GAAGCTCATATTGTATCTTTTTTAAAAACCATGTTGTGGAAAAAAGCCAGAGTGACAAGTG
ACAAAATCTATTTAGGAACTCTGTGTATGAATCCTGATTTTAACTGCTAGGATTCAGCTA
AATTTCTGAGCTTTATGATCTGTGGAAATTTGGAATGAAATCGAATTCATTTTGTACATA
CATAGTATATTAAAACTATATAATAGTTCATAGAAATGTTCAGTAATGAAAAAATATATC
CAATCAGAGCCATCCCGAAAAAAAAAAAAAAA (SEQ ID No.: 8827)

The FASTA file, including the sequence of NM_002310, was masked using the RepeatMasker web interface (Smit, AFA & Green, P RepeatMasker at http://ftp.genome.washington.edu/RM/RepeatMasker.html, Smit and Green). Specifically, during masking, the following types of sequences were replaced with "N's": SINE/MIR & LINE/L2, LINE/L1 , LTR/MaLR, LTR/Retroviral , Alu, and other low

informational content sequences such as simple repeats. Below is the sequence following masking:

CTCTCTCCCAGAACGTGTCTCTGCTGCAAGGCACCGGGCCCTTTCGCTCTGCAGAACTG
CACTTGCAAGACCATTATCAACTCCTAATCCCAGCTCAGAAAGGGAGCCTCTGCGACTC
ATTCATCGCCCTCCAGGACTGACTGCATTGCACAGATGATGGATATTTACGTATGTTTG
AAACGACCATCCTGGATGGTGGACAATAAAAGAATGAGGACTGCTTCAAATTTCCAGTG
GCTGTTATCAACATTTATTCTTCTATATCTAATGAATCAAGTAAATAGCCAGAAAAAGG
GGGCTCCTCATGATTTGAAGTGTGTAACTAACAATTTGCAAGTGTGGAACTGTTCTTGG
AAAGCACCCTCTGGAACAGGCCGTGGTACTGATTATGAAGTTTGCATTGAAAACAGGTC
CCGTTCTTGTTATCAGTTGGAGAAAACCAGTATTAAAATTCCAGCTCTTTCACATGGTG
ATTATGAAATAACAATAAATTCTCTACATGATTTTGGAAGTTCTACAAGTAAATTCACA
CTAAATGAACAAAACGTTTCCTTAATTCCAGATACTCCAGAGATCTTGAATTTGTCTGC
TGATTTCTCAACCTCTACATTATACCTAAAGTGGAACGACAGGGGTTCAGTTTTTCCAC
ACCGCTCAAATGTTATCTGGGAAATTAAAGTTCTACGTAAAGAGAGTATGGAGCTCGTA
AAATTAGTGACCCACAACACAACTCTGAATGGCAAAGATACACTTCATCACTGGAGTTG
GGCCTCAGATATGCCCTTGGAATGTGCCATTCATTTTGTGGAAATTAGATGCTACATTG
ACAATCTTCATTTTTCTGGTCTCGAAGAGTGGAGTGACTGGAGCCCTGTGAAGAACATT
TCTTGGATACCTGATTCTCAGACTAAGGTTTTTCCTCAAGATAAAGTGATACTTGTAGG
CTCAGACATAACATTTTGTTGTGTGAGTCAAGAAAAAGTGTTATCAGCACTGATTGGCC
ATACAAACTGCCCCTTGATCCATCTTGATGGGGAAAATGTTGCAATCAAGATTCGTAAT
ATTTCTGTTTCTGCAAGTAGTGGAACAAATGTAGTTTTTACAACCGAAGATAACATATT
TGGAACCGTTATTTTTGCTGGATATCCACCAGATACTCCTCAACAACTGAATTGTGAGA
CACATGATTTAAAAGAAATTATATGTAGTTGGAATCCAGGAAGGGTGACAGCGTTGGTG
GGCCCACGTGCTACAAGCTACACTTTAGTTGAAAGTTTTTCAGGAAAATATGTTAGACT
TAAAAGAGCTGAAGCACCTACAAACGAAAGCTATCAATTATTATTTCAAATGCTTCCAA
ATCAAGAAATATATAATTTTACTTTGAATGCTCACAATCCGCTGGGTCGATCACAATCA
ACAATTTTAGTTAATATAACTGAAAAAGTTTATCCCCATACTCCTACTTCATTCAAAGT
GAAGGATATTAATTCAACAGCTGTTAAACTTTCTTGGCATTTACCAGGCAACTTTGCAA
AGATTAATTTTTTATGTGAAATTGAAATTAAGAAATCTAATTCAGTACAAGAGCAGCGG
AATGTCACAATCAAAGGAGTAGAAAATTCAAGTTATCTTGTTGCTCTGGACAAGTTAAA
TCCATACACTCTATATACTTTTCGGATTCGTTGTTCTACTGAAACTTTCTGGAAATGGA

GCAAATGGAGCAATAAAAAACAACATTTAACAACAGAAGCCAGTCCTTCAAAGGGGCCT
GATACTTGGAGAGAGTGGAGTTCTGATGGAAAAAATTTAATAATCTATTGGAAGCCTTT
ACCCATTAATGAAGCTAATGGAAAAATACTTTCCTACAATGTATCGTGTTCATCAGATG
AGGAAACACAGTCCCTTTCTGAAATCCCTGATCCTCAGCACAAAGCAGAGATACGACTT
GATAAGAATGACTACATCATCAGCGTAGTGGCTAAAAATTCTGTGGGCTCATCACCACC
TTCCAAAATAGCGAGTATGGAAATTCCAAATGATGATCTCAAAATAGAACAAGTTGTTG
GGATGGGAAAGGGGATTCTCCTCACCTGGCATTACGACCCCAACATGACTTGCGACTAC
GTCATTAAGTGGTGTAACTCGTCTCGGTCGGAACCATGCCTTATGGACTGGAGAAAAGT
TCCCTCAAACAGCACTGAAACTGTAATAGAATCTGATGAGTTTCGACCAGGTATAAGAT
ATAATTTTTTCCTGTATGGATGCAGAAATCAAGGATATCAATTATTACGCTCCATGATT
GGATATATAGAAGAATTGGCTCCCATTGTTGCACCAAATTTTACTGTTGAGGATACTTC
TGCAGATTCGATATTAGTAAAATGGGAAGACATTCCTGTGGAAGAACTTAGAGGCTTTT
TAAGAGGATATTTGTTTTACTTTGGAAAAGGAGAAAGAGACACATCTAAGATGAGGGTT
TTAGAATCAGGTCGTTCTGACATAAAAGTTAAGAATATTACTGACATATCCCAGAAGAC
ACTGAGAATTGCTGATCTTCAAGGTAAAACAAGTTACCACCTGGTCTTGCGAGCCTATA
CAGATGGTGGAGTGGGCCCGGAGAAGAGTATGTATGTGGTGACAAAGGAAAATTCTGTG
GGATTAATTATTGCCATTCTCATCCCAGTGGCAGTGGCTGTCATTGTTGGAGTGGTGAC
AAGTATCCTTTGCTATCGGAAACGAGAATGGATTAAAGAAACCTTCTACCCTGATATTC
CAAATCCAGAAAACTGTAAAGCATTACAGTTTCAAAAGAGTGTCTGTGAGGGAAGCAGT
GCTCTTAAAACATTGGAAATGAATCCTTGTACCCCAAATAATGTTGAGGTTCTGGAAAC
TCGATCAGCATTTCCTAAAATAGAAGATACAGAAATAATTTCCCCAGTAGCTGAGCGTC
CTGAAGATCGCTCTGATGCAGAGCCTGAAAACCATGTGGTTGTGTCCTATTGTCCACCC
ATCATTGAGGAAGAAATACCAAACCCAGCCGCAGATGAAGCTGGAGGGACTGCACAGGT
TATTTACATTGATGTTCAGTCGATGTATCAGCCTCAAGCAAAACCAGAAGAAGAACAAG
AAAATGACCCTGTAGGAGGGGCAGGCTATAAGCCACAGATGCACCTCCCCATTAATTCT
ACTGTGGAAGATATAGCTGCAGAAGAGGACTTAGATAAAACTGCGGGTTACAGACCTCA
GGCCAATGTAAATACATGGAATTTAGTGTCTCCAGACTCTCCTAGATCCATAGACAGCA
ACAGTGAGATTGTCTCATTTGGAAGTCCATGCTCCATTAATTCCCGACAATTTTTGATT
CCTCCTAAAGATGAAGACTCTCCTAAATCTAATGGAGGAGGGTGGTCCTTTACAAACTT
TTTTCAGAACAAACCAAACGATTAACAGTGTCACCGTGTCACTTCAGTCAGCCATCTCA
ATAAGCTCTTACTGCTAGTGTTGCTACATCAGCACTGGGCATTCTTGGAGGGATCCTGT
GAAGTATTGTTAGGAGGTGAACTTCACTACATGTTAAGTTACACTGAAAGTTCATGTGC

TTTTAATGTAGTCTAAAAGCCAAAGTATAGTGACTCAGAATCCTCAATCCACAAAACTC
AAGATTGGGAGCTCTTTGTGATCAAGCCAAAGAATTCTCATGTACTCTACCTTCAAGAA
GCATTTCAAGGCTAATACCTACTTGTACGTACATGTAAAACAAATCCCGCCGCAACTGT
TTTCTGTTCTGTTGTTTGTGGTTTTCTCATATGTATACTTGGTGGAATTGTAAGTGGAT
TTGCAGGCCAGGGAGAAAATGTCCAAGTAACAGGTGAAGTTTATTTGCCTGACGTTTAC
TCCTTTCTAGATGAAAACCAAGCACAGATTTTAAAACTTCTAAGATTATTCTCCTCTAT
CCACAGCATTCACNNNNNNNNNNNNNNNNNNNNNNGTAGTGACAGCGATTTAGTGTTTT
GTTTGATAAAGTATGCTTATTTCTGTGCCTACTGTATAATGGTTATCAAACAGTTGTCT
CAGGGGTACAAACTTTGAAAACAAGTGTGACACTGACCAGCCCAAATCATAATCATGTT
TTCTTGCTGTGATAGGTTTTGCTTGCCTTTTCATTATTTTTTAGCTTTTATGCTTGCTT
CCATTATTTCAGTTGGTTGCCCTAATATTTAAAATTTACACTTCTAAGACTAGAGACCC
ACATTTTTTAAAAATCATTTTATTTTGTGATACAGTGACAGCTTTATATGAGCAAATTC
AATATTATTCATAAGCATGTAATTCCAGTGACTTACTATGTGAGATGACTACTAAGCAA
TATCTAGCAGCGTTAGTTCCATATAGTTCTGATTGGATTTCGTTCCTCCTGAGGAGACC
ATGCCGTTGAGCTTGGCTACCCAGGCAGTGGTGATCTTTGACACCTTCTGGTGGATGTT
CCTCCCACTCATGAGTCTTTTCATCATGCCACATTATCTGATCCAGTCCTCACATTTTT
AAATATAAAACTAAAGAGAGAATGCTTCTTACAGGAACAGTTACCCAAGGGCTGTTTCT
TAGTAACTGTCATAAACTGATCTGGATCCATGGGCATACCTGTGTTCGAGGTGCAGCAA
TTGCTTGGTGAGCTGTGCAGAATTGATTGCCTTCAGCACAGCATCCTCTGCCCACCCTT
GTTTCTCATAAGCGATGTCTGGAGTGATTGTGGTTCTTGGAAAAGCAGAAGGAAAAACT
AAAAAGTGTATCTTGTATTTTCCCTGCCCTCAGGTTGCCTATGTATTTTACCTTTTCAT
ATTTAAGGCAAAAGTACTTGAAAATTTTAAGTGTCCGAATAAGATATGTCTTTTTTGTT
TGTTTTTTTTGGTTGGTTGTTTGTTTTTTATCATCTGAGATTCTGTAATGTATTTGCAA
ATAATGGATCAATTAATTTTTTTTGAAGCTCATATTGTATCTTTTTAAAAACCATGTTG
TGGAAAAAAGCCAGAGTGACAAGTGACAAAATCTATTTAGGAACTCTGTGTATGAATCC
TGATTTTAACTGCTAGGATTCAGCTAAATTTCTGAGCTTTATGATCTGTGGAAATTTGG
AATGAAATCGAATTCATTTTGTACATACATAGTATATTAAAACTATATAATAGTTCATA
GAAATGTTCAGTAATGAAAAAATATATCCAATCAGAGCCATCCCGAAAAAAAAAAAAAA
A SEQ ID No.: 8828

The length of this sequence was determined using batch, automated computational methods and the sequence, as sense strand, its length, and the desired location of the probe sequence near the 3' end of the mRNA was submitted to Array Designer Ver 1.1 (Premier Biosoft International, Palo Alto, CA). Search quality was set at 100%, number of best probes set at 1, length range set at 50 base pairs, Target Tm set at 75 C. degrees plus or minus 5 degrees, Hairpin max deltaG at 6.0 -kcal/mol., Self dimmer max deltaG at 6.0 –kcal/mol, Run/repeat (dinucleotide) max length set at 5, and Probe site minimum overlap set at 1. When none of the 49 possible probes met the criteria, the probe site would be moved 50 base pairs closer to the 5' end of the sequence and resubmitted to Array Designer for analysis. When no possible probes met the criteria, the variation on melting temperature was raised to plus and minus 8 degrees and the number of identical basepairs in a run increased to 6 so that a probe sequence was produced.

In the sequence above, using the criteria noted above, Array Designer Ver 1.1 designed a probe corresponding to oligonucleotide number 2280 in Table 8 and is indicated by underlining in the sequence above. It has a melting temperature of 68.4 degrees Celsius and a max run of 6 nucleotides and represents one of the cases where the criteria for probe design in Array Designer Ver 1.1 were relaxed in order to obtain an oligonucleotide near the 3' end of the mRNA (Low melting temperature was allowed).

Clone 463D12

Clone 463D12 was sequenced and compared to the nr, dbEST, and UniGene databases at NCBI using the BLAST search tool. The sequence matched accession number AI184553, an EST sequence with the definition line "qd60a05.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1733840 3' similar to gb:M29550 PROTEIN PHOSPHATASE 2B CATALYTIC SUBUNIT 1 (HUMAN);, mRNA sequence." The E value of the alignment was 1.00×10^{-118}. The GenBank sequence begins with a poly-T region, suggesting that it is the antisense strand, read 5' to 3'. The beginning of this sequence is complementary to the 3' end of the mRNA sense strand. The accession number for this sequence was included in a text file of accession numbers representing antisense sequences. Sequences for antisense strand mRNAs were obtained by uploading a text file containing desired accession numbers as an Entrez

search query using the Batch Entrez web interface and saving the results locally as a FASTA file. The following sequence was obtained, and the region of alignment of clone 463D12 is outlined:

TTTTTTTTTTTTTTCTTAAATAGCATTTATTTTCTCTCAAAAAGCCTATTATGTACTAA
CAAGTGTTCCTCTAAATTAGAAAGGCATCACTACTAAAATTTTATACATATTTTTTATA
TAAGAGAAGGAATATTGGGTTACAATCTGAATTTCTCTTTATGATTTCTCTTAAAGTAT
AGAACAGCTATTAAAATGACTAATATTGCTAAAATGAAGGCTACTAAATTTCCCCAAGA
ATTTCGGTGGAATGCCCAAAAATGGTGTTAAGATATGCAGAAGGGCCCATTTCAAGCAA
AGCAATCTCTCCACCCCTTCATAAAAGATTTAAGCTAAAAAAAAAAAAAAAAAGAAGAAA
ATCCAACAGCTGAAGACATTGGGCTATTTATAAATCTTCTCCCAGTCCCCCAGACAGCC
TCACATGGGGGCTGTAAACAGCTAACTAAAATATCTTTGAGACTCTTATGTCCACACCC
ACTGACACAAGGAGAGCTGTAACCACAGTGAAACTAGACTTTGCTTTCCTTTAGCAAGT
ATGTGCCTATGATAGTAAACTGGAGTAAATGTAACAGTAATAAAACAAATTTTTTTTAA
AAATAAAAATTATACCTTTTTCTCCAACAAACGGTAAAGACCACGTGAAGACATCCATA
AAATTAGGCAACCAGTAAAGATGTGGAGAACCAGTAAACTGTCGAAATTCATCACATTA
TTTTCATACTTTAATACAGCAGCTTTAATTATTGGAGAACATCAAAGTAATTAGGTGCC
GAAAAACATTGTTATTAATGAAGGGAACCCCTGACGTTTGACCTTTTCTGTACCATCTA
TAGCCCTGGACTTGA (SEQ ID No.: 8829)

The FASTA file, including the sequence of AA184553, was then masked using the RepeatMasker web interface, as shown below. The region of alignment of clone 463D12 is outlined.

TTTTTTTTTTTTTTCTTAAATAGCATTTATTTTCTCTCAAAAAGCCTATTATGTACTAA
CAAGTGTTCCTCTAAATTAGAAAGGCATCACTACNNNNNNNNNNNNNNNNNNNNNNNN
NNNGAGAAGGAATATTGGGTTACAATCTGAATTTCTCTTTATGATTTCTCTTAAAGTAT
AGAACAGCTATTAAAATGACTAATATTGCTAAAATGAAGGCTACTAAATTTCCCCAAGA
ATTTCGGTGGAATGCCCAAAAATGGTGTTAAGATATGCAGAAGGGCCCATTTCAAGCAA
AGCAATCTCTCCACCCCTTCATAAAAGATTTAAGCTAAAAAAAAAAAAAAAAAGAAGAAA
ATCCAACAGCTGAAGACATTGGGCTATTTATAAATCTTCTCCCAGTCCCCCAGACAGCC
TCACATGGGGGCTGTAAACAGCTAACTAAAATATCTTTGAGACTCTTATGTCCACACCC

ACTGACACAAGGAGAGCTGTAACCACAGTGAAACTAGACTTTGCTTTCCTTTAGCAAGT
ATGTGCCTATGATAGTAAACTGGAGTAAATGTAACAGNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNCCTTTTTCTCCAACAAACGGTAAAGACCACGTGAAGACATCCATA
AAATTAGGCAACCAGTAAAGATGTGGAGAACCAGTAAACTGTCGAAATTCATCACATTA
TTTTCATACTTTAATACAGCAGCTTTAATTATTGGAGAACATCAAAGTAATTAGGTGCC
GAAAAACATTGTTATTAATGAAGGGAACCCCTGACGTTTGACCTTTTCTGTACCATCTA
TAGCCCTGGACTTGA Masked version of 463D12 sequence. (SEQ ID NO:8830)

The sequence was submitted to Array Designer as described above, however, the desired location of the probe was indicated at base pair 50 and if no probe met the criteria, moved in the 3' direction. The complementary sequence from Array Designer was used, because the original sequence was antisense. The oligonucleotide designed by Array Designer corresponds to oligonucleotide number 4342 in Table 8 and is complementary to the underlined sequence above. The probe has a melting temperature of 72.7 degrees centigrade and a max run of 4 nucleotides.

Clone 72D4

Clone 72D4 was sequenced and compared to the nr, dbEST, and UniGene databases at NCBI using the BLAST search tool. No significant matches were found in any of these databases. When compared to the human genome draft, significant alignments were found to three consecutive regions of the reference sequence NT_008060, as depicted below, suggesting that the insert contains three spliced exons of an unidentified gene.

Residue numbers on clone 72D4 sequence	Matching residue numbers on NT_008060
1 – 198	478646 – 478843
197 – 489	479876 – 480168
491 – 585	489271 – 489365

Because the reference sequence contains introns and may represent either the coding or noncoding strand for this gene, BioCardia's own sequence file was used to design the oligonucleotide. Two complementary probes were designed to ensure that the

sense strand was represented. The sequence of the insert in clone 72D4 is shown below, with the three putative exons outlined.

CAGGTCACACAGCACATCAGTGGCTACATGTGAGCTCAGACCTGGGTCTGCT
GCTGTCTGTCTTCCCAATATCCATGACCTTGACTGATGCAGGTGTCTAGGGAT
ACGTCCATCCCCGTCCTGCTGGAGCCCAGAGCACGGAAGCCTGGCCCTCCGA
GGAGACAGAAGGGAGTGTCGGACACCATGACGAGAGCTTGGCAGAATAAAT
AACTTCTTTAAACAATTTTACGGCATGAAGAAATCTGGACCAGTTTATTAAAT
GGGATTTCTGCCACAAACCTTGGAAGAATCACATCATCTTANNCCCAAGTGA
AAACTGTGTTGCGTAACAAAGAACATGACTGCGCTCCACACATACATCATTG
CCCGGCGAGGCGGGACACAAGTCAACGACGGAACACTTGAGACAGGCCTAC
AACTGTGCACGGGTCAGAAGCAAGTTTAAGCCATACTTGCTGCAGTGAGACT
ACATTTCTGTCTATAGAAGATACCTGACTTGATCTGTTTTTCAGCTCCAGTTC
CCAGATGTGCGTGTTGTGGTCCCCAAGTATCACCTTCCAATTTCTGGGAGCA
GTGCTCTGGCCGGATCCTTGCCGCGCGGATAAAAAC (SEQ ID NO.: 8445)

The sequence was submitted to RepeatMasker, but no repetitive sequences were found. The sequence shown above was used to design the two 50-mer probes using Array Designer as described above. The probes are shown in bold typeface in the sequence depicted below. The probe in the sequence is oligonucleotide number 6415 (SEQ ID NO.: 6415) in Table 8 and the complementary probe is oligonucleotide number 6805 (SEQ ID NO.:6805).

CAGGTCACACAGCACATCAGTGGCTACATGTGAGCTCAGACCTGGGTCTGCTGCTGTCT
GTCTTCCCAATATCCATGACCTTGACTGATGCAGGTGTCTAGGGATACGTCCATCCCCG
TCCTGCTGGAGCCCAGAGCACGGAAGCCTGGCCCTCCGAGGAGACAGAAGGGAGTGTCG
GACACCATGACGAGAGCTTGGCAGAATAAATAACTTCTTTAAACAATTTTACGGCATGA
AGAAATCTGGACCAGTTTATTAAATGGGATTTCTGCCACAAACCTTGGAAGAATCACAT
CATCTTANNCCCAAGTGAAAACTGTGTTGCGTAACAAAGAACATGACTGCGCTCCACAC
ATACATCATTGCCCGGCGAGGCGGGACACAAGTCAACGACGGAACACTTGAGACAGGCC

TACAACTGTGCACGGGTCAGAAGCAAGTTTAAGCCATACTTGCTGCAGTGAGACTACAT

TTCTGTCTATAGAAGATACCTGACTTGATCTGTTTTTCAGCTC**CAGTTCCCAGATGTGC**

←----GTCAAGGGTCTACACG

GTGTTGTGGTCCCCAAGTATCACCTTCCAATTTCTGGGAG--→

CACAACACCAGGGGTTCATAGTGGAAGGTTAAAG-5'

CAGTGCTCTGGCCGGATCCTTGCCGCGCGGATAAAAACT---→

Confirmation of probe sequence

Following probe design, each probe sequence was confirmed by comparing the sequence against dbEST, the UniGene cluster set, and the assembled human genome using BLASTn at NCBI. Alignments, accession numbers, gi numbers, UniGene cluster numbers and names were examined and the most common sequence used for the probe. The final probe set was compiled into Table 8.

Example 22 - Production of an array of 8000 spotted 50mer oligonucleotides

We produced an array of 8000 spotted 50mer oligonucleotides. Examples 20 and 21 exemplify the design and selection of probes for this array.

Sigma-Genosys (The Woodlands, TX) synthesized un-modified 50-mer oligonucleotides using standard phosphoramidite chemistry, with a starting scale of synthesis of 0.05 μmole (see, e.g., R. Meyers, ed. (1995) Molecular Biology and Biotechnology: A Comprehensive Desk Reference). Briefly, to begin synthesis, a 3' hydroxyl nucleoside with a dimethoxytrityl (DMT) group at the 5' end was attached to a solid support. The DMT group was removed with trichloroacetic acid (TCA) in order to free the 5'-hydroxyl for the coupling reaction. Next, tetrazole and a phosphoramidite derivative of the next nucleotide were added. The tetrazole protonates the nitrogen of the phosphoramidite, making it susceptible to nucleophilic attack. The DMT group at the 5'-end of the hydroxyl group blocks further addition of nucleotides in excess. Next, the inter-nucleotide linkage was converted to a phosphotriester bond in an oxidation step using an oxidizing agent and water as the oxygen donor. Excess nucleotides were filtered

out and the cycle for the next nucleotide was started by the removal of the DMT protecting group. Following the synthesis, the oligo was cleaved from the solid support. The oligonucleotides were desalted, resuspended in water at a concentration of 100 or 200 μM, and placed in 96-deep well format. The oligonucleotides were re-arrayed into Whatman Uniplate 384-well polyproylene V bottom plates. The oligonucleotides were diluted to a final concentration 30 μM in 1X Micro Spotting Solution Plus (Telechem/arrayit.com, Sunnyvale, CA) in a total volume of 15 μl. In total, 8,031 oligonucleotides were arrayed into twenty-one 384-well plates.

Arrays were produced on Telechem/arrayit.com Super amine glass substrates (Telechem/arrayit.com), which were manufactured in 0.1 mm filtered clean room with exact dimensions of 25x76x0.96 mm. The arrays were printed using the Virtek Chipwriter with a Telechem 48 pin Micro Spotting Printhead. The Printhead was loaded with 48 Stealth SMP3B TeleChem Micro Spotting Pins, which were used to print oligonucleotides onto the slide with the spot size being 110-115 microns in diameter.

Example 23- Amplification, labeling, and hybridization of total RNA to an oligonucleotide microarray

Amplification, labeling, hybridization and scanning

Samples consisting of at least 2 μg of intact total RNA were further processed for array hybridization. Amplification and labeling of total RNA samples was performed in three successive enzymatic reactions. First, a single-stranded DNA copy of the RNA was made (hereinafter, "ss-cDNA"). Second, the ss-cDNA was used as a template for the complementary DNA strand, producing double-stranded cDNA (hereinafter, "ds-cDNA, or cDNA"). Third, linear amplification was performed by in vitro transcription from a bacterial T_7 promoter. During this step, fluorescent-conjugated nucleotides were incorporated into the amplified RNA (hereinafter, "aRNA").

The first strand cDNA was produced using the Invitrogen kit (Superscript II). The first strand cDNA was produced in a reaction composed of 50 mM Tris-HCl (pH 8.3), 75 mM KCl, and 3 mM $MgCl_2$ (1x First Strand Buffer, Invitrogen), 0.5 mM dGTP, 0.5 mM dATP, 0.5 mM dTTP, 0.5 mM dCTP, 10 mM DTT, 10 U reverse transcriptase (Superscript II, Invitrogen, #18064014), 15 U RNase inhibitor (RNAGuard, Amersham

Pharmacia, #27-0815-01), 5 μM T7T24 primer (5′-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGGTTTTTTTTTTTT TTTTTTTTTTTT-3′), (SEQ ID NO.:8831) and 2 μg of selected sample total RNA. Several purified, recombinant control mRNAs from the plant Arabidopsis thaliana were added to the reaction mixture: 20 pg of CAB and RCA, 14 pg of LTP4 and NAC1, and 2 pg of RCP1 and XCP2 (Stratagene, #252201, #252202, #252204, #252208, #252207, #252206 respectively). The control RNAs allow the estimate of copy numbers for individual mRNAs in the clinical sample because corresponding sense oligonucleotide probes for each of these plant genes are present on the microarray. The final reaction volume of 40 μl was incubated at 42°C for 60 min.

For synthesis of the second cDNA strand, DNA polymerase and RNase were added to the previous reaction, bringing the final volume to 150 μl. The previous contents were diluted and new substrates were added to a final concentration of 20 mM Tris-HCl (pH 7.0) (Fisher Scientific, Pittsburgh, PA #BP1756-100), 90 mMKCl (Teknova, Half Moon Bay, CA, #0313-500) , 4.6 mM $MgCl_2$ (Teknova, Half Moon Bay, CA, #0304-500), 10 mM$(NH_4)_2SO_4$ (Fisher Scientific #A702-500)(1x Second Strand buffer, Invitrogen), 0.266 mM dGTP, 0.266 mM dATP, 0.266 mM dTTP, 0.266 mM dCTP, 40 U *E. coli* DNA polymerase (Invitrogen, #18010-025), and 2 U RNaseH (Invitrogen, #18021-014). The second strand synthesis took place at 16°C for 120 minutes.

Following second-strand synthesis, the ds-cDNA was purified from the enzymes, dNTPs, and buffers before proceeding to amplification, using phenol-chloroform extraction followed by ethanol precipitation of the cDNA in the presence of glycogen. Alternatively, a silica-gel column is used to purify the cDNA (e.g. Qiaquick PCR cleanup from Qiagen, #28104). The cDNA was collected by centrifugation at >10,000 ×g for 30 minutes, the supernatant is aspirated, and 150 μl of 70% ethanol, 30% water was added to wash the DNA pellet. Following centrifugation, the supernatant was removed, and residual ethanol was evaporated at room temperature.

Linear amplification of the cDNA was performed by in vitro transcription of the cDNA. The cDNA pellet from the step described above was resuspended in 7.4 μl of water, and in vitro transcription reaction buffer was added to a final volume of 20 μl

containing 7.5 mM GTP, 7.5 mM ATP, 7.5 mM TTP, 2.25 mM CTP, 1.025 mM Cy3-conjugated CTP (Perkin Elmer; Boston, MA, #NEL-580), 1x reaction buffer (Ambion, Megascript Kit, Austin, TX and #1334) and 1 % T_7 polymerase enzyme mix (Ambion, Megascript Kit, Austin, TX and #1334). This reaction was incubated at 37°C overnight. Following in vitro transcription, the RNA was purified from the enzyme, buffers, and excess NTPs using the RNeasy kit from Qiagen (Valencia, CA; # 74106) as described in the vendor's protocol. A second elution step was performed and the two eluates were combined for a final volume of 60 µl. RNA is quantified using an Agilent 2100 bioanalyzer with the RNA 6000 nano LabChip.

Reference RNA was prepared as described above, except that 10 µg of total RNA was the starting material for amplification, and Cy5-CTP was incorporated instead of Cy3CTP. Reference RNA from five reactions was pooled together and quantitated as described above.

Hybridization to an array

RNA was prepared for hybridization as follows: for an 18mm×55mm array, 20 µg of amplified RNA (aRNA) was combined with 20 µg of reference aRNA. The combined sample and reference aRNA was concentrated by evaporating the water to 5 µl in a vacuum evaporator. Five µl of 20 mM zinc acetate was added to the aRNA and the mix incubated at 60°C for 10 minutes to fragment the RNA into 50-200 bp pieces. Following the incubation, 40 µl of hybridization buffer was added to achieve final concentrations of 5×SSC and 0.20 %SDS with 0.1 µg/ul of Cot-1 DNA (Invitrogen) as a competitor DNA. The final hybridization mix was heated to 98°C, and then reduced to 50°C at 0.1°C per second.

Alternatively, formamide is included in the hybridization mixture to lower the hybridization temperature.

The hybridization mixture was applied to the microarray surface, covered with a glass coverslip (Corning, #2935-246), and incubated in a humidified chamber (Telechem, AHC-10) at 62°C overnight. Following incubation, the slides were washed in 2×SSC, 0.1% SDS for two minutes, then in 2×SSC for two minutes, then in 0.2×SSC for two

minutes. The arrays were spun at 1000×g for 2 minutes to dry them. The dry microarrays are then scanned by methods described above.

Example 24: Analysis of Human Transplant Patient Mononuclear cell RNA Hybridized to a 24,000 Feature Microarray.

Patients who had recently undergone cardiac transplant and were being monitored for rejection by biopsy were selected and enrolled in a clinical study, as described in Example 11. Blood was drawn from several patients and mononuclear cells isolated as described in Example 8. The rejection grade determined from the biopsy is presented in Table **9** for some of the patient samples. Four samples (14-0001-2, 14-0001-3, 14-0005-1 and 14-0005-2) from one center were selected for further examination. Two sets of paired samples were available that allowed comparison of severe rejection (rejection grade 3A) to minimal or no rejection (rejection grade 1 or 0). These two groups are designated "high rejection grade" and "low rejection grade", respectively.

Additional RNA was isolated from the mononuclear cells of enrolled cardiac allograft recipients as described in Example 8. The yield of RNA from 8 ml of blood is shown in Table **9**, below.

1 or 2 μg of total RNA was amplified by making cDNA copies using a T7T24 primer and subsequent in vitro transcription, as described in Example 23. This "target" amplified RNA was labeled by incorporation of Cy3-conjugated nucleotides, as described in Example 23. The amplified RNA was quantified by analysis at A260 on a spectrophotometer.

Hybridization to the 8,000 probe (24,000-feature) microarray (described in Examples 20-22) was performed essentially as described in Example 23. 20 μg of amplified and labeled RNA was combined with 20 μg of R50 reference RNA that was labeled and prepared as described in Example 9.

The sample and reference amplified and labeled RNAs were combined and fragmented at 95°C for 30 min, as described in Example 23. The fragmented RNA was mixed with 40 μl of hybridization solution (to bring the total to 50 μl) and applied to the 8,000-probe, 24,000-feature microarray and covered with a 21mm×60mm coverslip. The arrays were hybridized overnight and washed as described in Example 23.

Once hybridized and washed, the arrays were scanned as described in Example 23. The full image produced by the Agilent scanner G2565AA was flipped, rotated, and split into two images (one for each signal channel) using TIFFSplitter (Agilent, Palo Alto, CA). The two channels are the output at 532 nm (Cy3-labeled sample) and 633 nm(Cy5-labeled R50). The individual images were loaded into GenePix 3.0 (Axon Instruments, Union City, CA) and the software was used to determine the median pixel intensity for each feature (F_i) and the median pixel intensity of the local background for each feature (B_i) in both channels. The standard deviation ($SDF_{i\ and}\ SDB_i$) for each is also determined. Features for which GenePix could not discriminate the feature from the background were "flagged", and the data were deleted from further consideration.

From the remaining data, the following calculations were performed.

The first calculation performed was the signal to noise ratio:

$$S/N = \frac{F_i - B_i}{SDB_i}$$

All features with a S/N less than 3 in either channel were removed from further consideration. All features that did not have GenePix flags and passed the S/N test were considered usable features. The background-subtracted signal (hereinafter, "BGSS") was calculated for each usable feature in each channel ($BGSS_i=F_i-B_i$).

The BGSS was used for the scaling step within each channel. The median BGSS for all usable features was calculated. The $BGSS_i$ for each feature was divided by the median BGSS. The median BGSS for the scaled data then became 1 for each channel on each array. This operation did not change the distribution of the data, but did allow each to be directly compared

The scaled $BGSS_i$ (S_i) for each feature was used to calculate the ratio of the Cy3 to the Cy5 signal:

$$R_n = \frac{Cy3S_i}{Cy5S_i}$$

The ratio data from the triplicate features were combined for each probe on the array. If all three features were still usable, their average was taken (R_p) and the coefficient of variation (hereinafter "CV") was determined. If the CV was less than 15%, the average was carried forward for that probe. If the CV was greater than 15% for the triplicate features, then the average of the two features with the closest R_n values were used. If there were only two usable features for a given probe, the average of the two features was used. If there was only one usable feature for a given probe, the value of that feature was used.

The logarithm of the average ratio was taken for each probe ($\log R_p$). This value was used for comparison among arrays. For comparison of gene expression in high rejection grade patients to gene expression from low rejection grade patients, the average was taken for each probe for hybridizations 107739 and 107741 (high rejection grades) and 107740 and 107742 (low rejection grades). Since there were only two patients, each with a change from high to low rejection grade, there should be less variability in the data than if all four samples were from different patients. The results of this comparison were plotted in Figure **9**. The X-axis is the high rejection grade average (the average of each probe for hybridizations of samples from high rejection grade patients) and the Y-axis is the low rejection grade average. There was complete data for 5562 probes, all plotted in Figure 9. Each "point" in the graph corresponded to a probe on the microarray.

A "cluster" of points were shaded in white. Points within the cluster represented genes with expression that is not significantly changed from one sample group to the other. The far ends of the cluster corresponded to genes that are expressed at either low or high levels in each group.

Outlier points, corresponding to genes with differential expression between high and low rejection grade patients, were shaded black and are further described in Table **10**. There was one point above the cluster (indicating that expression was relatively higher in the low rejection grade than in the low rejection grade), and 7 points below the cluster (indicating that expression was relatively higher in the high rejection grade than in the low rejection grade).

Many of the differentially expressed genes had unknown or poorly described functions. One, corresponding to probe number 8091, was known in the public databases only as a predicted mRNA and protein.

Using the data from samples 107739 (Grade 3A rejection) and 107742 (Grade 0), a scaled ratio of sample (Cy3) to reference (Cy5) expression was determined using the same techniques. The ratio of was taken of these scaled ratios, denoted "the ratio of scaled ratios (hereinafter, "SR"). Replicate features were not combined and all probes with S/N < 3 in either channel were filtered out. Some probes with differential expression between these two samples are shown in Figure **10**. In this Figure, the probes are sorted from the top to the bottom by relative expression in the first grade 0 sample vs grade 3A (ratio of SRs, grade 0/3A).

Diagnostic accuracy for sample classification is determined using additional samples and suitable methods for correlation analysis.

Comparing Figure 10 and Table 10, genes of particular interest include those corresponding to SEQ ID NO:2476, SEQ ID NO: 2407, SEQ ID NO:2192, SEQ ID NO: 2283, SEQ ID NO:6025, SEQ ID NO: 4481, SEQ ID NO:3761, SEQ ID NO: 3791, SEQ ID NO:4476, SEQ ID NO: 4398, SEQ ID NO:7401, SEQ ID NO: 1796, SEQ ID NO:4423, SEQ ID NO: 4429, SEQ ID NO:4430, SEQ ID NO: 4767, SEQ ID NO:4829 and SEQ ID NO: 8091.

Table 1

Disease Classification	Disease/Patient Group
Cardiovascular Disease	Atherosclerosis
	Unstable angina
	Myocardial Infarction
	Restenosis after angioplasty
	Congestive Heart Failure
	Myocarditis
	Endocarditis
	Endothelial Dysfunction
	Cardiomyopathy
	Cardiovascular drug use
Endocrine Disease	Diabetes Mellitus I and II
	Thyroiditis
	Addisson's Disease
Infectious Disease	Hepatitis A, B, C, D, E, G
	Malaria
	Tuberculosis
	HIV
	Pneumocystis Carinii
	Giardia
	Toxoplasmosis
	Lyme Disease
	Rocky Mountain Spotted Fever
	Cytomegalovirus
	Epstein Barr Virus
	Herpes Simplex Virus
	Clostridium Dificile Colitis
	Meningitis (all organisms)
	Pneumonia (all organisms)
	Urinary Tract Infection (all organisms)
	Infectious Diarrhea (all organisms)
	Anti-infectious drug use
Angiogenesis	Pathologic angiogenesis
	Physiologic angiogenesis
	Treatment induced angiogenesis
	Pro or anti-angiogenic drug use
Inflammatory/Rheumatic	Rheumatoid Arthritis
	Systemic Lupus Erythematosis
	Sjogrens Disease
	CREST syndrome
	Scleroderma
	Ankylosing Spondylitis
	Crohn's
	Ulcerative Colitis
	Primary Sclerosing Cholangitis

Table 1 (continued)

Disease Classification	Disease/Patient Group
Inflammatory/Rheumatic	Appendicitis
	Diverticulitis
	Primary Biliary Sclerosis
	Wegener's Granulomatosis
	Polyarteritis nodosa
	Whipple's Disease
	Psoriasis
	Microscopic Polyanngiitis
	Takayasu's Disease
	Kawasaki's Disease
	Autoimmune hepatitis
	Asthma
	Churg-Strauss Disease
	Beurger's Disease
	Raynaud's Disease
	Cholecystitis
	Sarcoidosis
	Asbestosis
	Pneumoconioses
	Antinflammatory drug use
Transplant Rejection	Heart
	Lung
	Liver
	Pancreas
	Bowel
	Bone Marrow
	Stem Cell
	Graft versus host disease
	Transplant vasculopathy
	Skin
	Cornea
	Immunosupressive drug use
Malignant Disorders	Leukemia
	Lymphoma
	Carcinoma
	Sarcoma
Neurological Disease	Alzheimer's Dementia
	Pick's Disease
	Multiple Sclerosis
	Guillain Barre Syndrome
	Peripheral Neuropathy

Table 2: Candidate genes, Database mining

Unigene clusters are listed.
Cluster numbers are defined as in Unigene build #133 uploaded on: Fri Apr 20 2001

CD50	Hs.99995
CD70 = CD27L	Hs.99899
MDC	Hs.97203
CD3z	Hs.97087
CD19	Hs.96023
	Hs.95388
CD3d	Hs.95327
	Hs.9456
interleukin 6	Hs.93913
phospholipaseA2	Hs.93304
Human mRNA for KIAA0128 gene, partial cds.	Hs.90998
CD48	Hs.901
heat shock 70kD protein 1A	Hs.8997
TxA2 receptor	Hs.89887
fragile X mental retardation protein (FMR-1)	Hs.89764
CD20	Hs.89751
ENA-78	Hs.89714
IL-2	Hs.89679
CD79b	Hs.89575
CD2	Hs.89476
SDF-1=CXCR4	Hs.89414
CD61	Hs.87149
IFN-g	Hs.856
CD34	Hs.85289
CD104	Hs.85266
CD8	Hs.85258
IGF-1	Hs.85112
CD103	Hs.851
IL-13	Hs.845
RPA1	Hs.84318
CD74	Hs.84298
CD132	Hs.84
CD18	Hs.83968
Cathepsin K	Hs.83942
CD80	Hs.838
CD46	Hs.83532
NFKB1	Hs.83428
IL-18	Hs.83077
interleukin 14	Hs.83004
L-selectin = CD62L	Hs.82848
CD107b	Hs.8262
CD69	Hs.82401
CD95	Hs.82359
CD53	Hs.82212

Homo sapiens cAMP responsive element binding protein 1 (CREB1) mRNA.	Hs.79194
Nucleolin (NCL)	Hs.79110
MAPK14	Hs.79107
CD100	Hs.79089
OX-2	Hs.79015
PCNA	Hs.78996
	Hs.78909
GRO-a	Hs.789
CDw32A	Hs.78864
H.sapiens mRNA for herpesvirus associated ubiquitin-specific protease (HAUSP).	Hs.78683
CD41b = LIBS1	Hs.785
ANXA1 (LPC1)	Hs.78225
CD31	Hs.78146
Homo sapiens TERF1 (TRF1)-interacting nuclear factor 2 (TINF2), mRNA.	Hs.7797
major histocompatibility complex, class I, B	Hs.77961
LOX1	Hs.77729
major histocompatibility complex, class II, DM alpha	Hs.77522
CD64	Hs.77424
CD71	Hs.77356
	Hs.77054
HLA-DRA	Hs.76807
CD105	Hs.76753
	Hs.76691
TNF-alpha	Hs.76507
LCP1	Hs.76506
TMSB4X	Hs.75968
PAI2	Hs.75716
MIP-1b	Hs.75703
CD58	Hs.75626
CD36	Hs.75613
hnRNP A2 / hnRNP B1	Hs.75598
CD124	Hs.75545
MIP-3a	Hs.75498
beta-2-microglobulin	Hs.75415
FPR1	Hs.753
Topo2B	Hs.75248
interleukin enhancer binding factor 2, 45kD	Hs.75117
chloride intracellular channel 1	Hs.74276
EGR3	Hs.74088
MIP-1a	Hs.73817
CD62P = p-selectin	Hs.73800
CD21	Hs.73792
APE	Hs.73722
IL12Rb2	Hs.73165

Table 2: Candidate genes, Database mining

Human lymphocyte specific interferon regulatory factor/interferon regulatory factor 4 (LSIRF/IRF4) mRNA, complete cds.	Hs.82132
IL-16	Hs.82127
DUT	Hs.82113
CDw121a	Hs.82112
PAI-1	Hs.82085
TGF-bR2	Hs.82028
CD117	Hs.81665
HLA-DPB1	Hs.814
NFKBIA	Hs.81328
CD6	Hs.81226
IL-1 RA	Hs.81134
UBE2B (RAD6B)	Hs.811
Lyn	Hs.80887
STAT4	Hs.80642
UBE2A (RAD6A)	Hs.80612
Fractalkine	Hs.80420
IK cytokine, down-regulator of HLA II	Hs.8024
	Hs.79933
CD79a	Hs.79630
	Hs.7942
nuclear factor, interleukin 3 regulated	Hs.79334
CD83	Hs.79197
DC-CK1	Hs.16530
CCR7	Hs.1652
TLR4	Hs.159239
EST	Hs.158975
EST	Hs.158966
EST	Hs.158965
EST	Hs.158943
EST	Hs.158894
EST	Hs.158877
EST	Hs.157815
EST	Hs.157813
ESTs	Hs.157569
immunoglobulin kappa constant	Hs.156110
INPP5D	Hs.155939
C3AR1	Hs.155935
PRKDC	Hs.155637

NFKB2	Hs.73090
I-309	Hs.72918
immunoglobulin superfamily, member 4	Hs.70337
IL-3	Hs.694
	Hs.6895
NTH1	Hs.66196
CD40L	Hs.652
IL-11R	Hs.64310
Homo sapiens toll-like receptor 2 (TLR2) mRNA.	Hs.63668
ferritin H chain	Hs.62954
IL8	Hs.624
Tissue Factor	Hs.62192
F-box only protein 7	Hs.5912
CD5	Hs.58685
guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1	Hs.5662
SCYA11	Hs.54460
IK1	Hs.54452
CCR1	Hs.516
Homo sapiens TRAIL receptor 2 mRNA, complete cds.	Hs.51233
CD11c	Hs.51077
CD66a	Hs.50964
JAK1	Hs.50651
Homo sapiens programmed cell death 4 (PDCD4), mRNA.	Hs.100407
SCYB13 (CXCL13)	Hs.100431
SMAD7	Hs.100602
RAD51L1 (RAD51B)	Hs.100669
PPARG	Hs.100724
transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)	Hs.101047
major histocompatibility complex, class I-like sequence	Hs.101840
immunoglobulin superfamily containing leucine-rich repeat	Hs.102171
CD166	Hs.10247
fibroblast tropomyosin TM30 (pl)	Hs.102824
interleukin 1 receptor-like 2	Hs.102865
GTF2H4	Hs.102910
	Hs.10326
Human ITAC (IBICK)	Hs.103982
novel protein with MAM domain	Hs.104311
ESTs, Weakly similar to interleukin enhancer binding factor 2 [H.sapiens]	Hs.105125

Table 2: Candidate genes, Database mining

MHC class II HLA-DRw53-associated glycoprotein	Hs.155122
CD73	Hs.153952
CD37	Hs.153053
IFNAR1	Hs.1513
Homo sapiens solute carrier family 21 (organic anion transporter), member 11 (SLC21A11), mRNA.	Hs.14805
EST	Hs.146627
SET translocation (myeloid leukemia-associated)	Hs.145279
EST	Hs.144119
ESTs	Hs.143534
STAT3	Hs.142258
CD96	Hs.142023
CD23	Hs.1416
EGR2	Hs.1395
CDw84	Hs.137548
CD55	Hs.1369
EST	Hs.135339
GM-CSF	Hs.1349
EST	Hs.133175
CD1a	Hs.1309
CD10	Hs.1298
HVEM	Hs.129708
C9	Hs.1290
C6	Hs.1282
C1R	Hs.1279
IL-1b	Hs.126256
CD9	Hs.1244
	Hs.12305
Homo sapiens Vanin 2 (VNN2) mRNA.	Hs.121102
Hsp10	Hs.1197
CD59	Hs.119663
CD51	Hs.118512
CD49a	Hs.116774
CD72	Hs.116481
HLA-DMB	Hs.1162
MCP-4	Hs.11383
	Hs.111554

Homo sapiens clone 24686 mRNA sequence.	Hs.105509
	Hs.105532
Homo sapiens granulysin (GNLY), transcript variant 519, mRNA.	Hs.105806
CD77	Hs.105956
RD RNA-binding protein	Hs.106061
	Hs.106673
	Hs.10669
Homo sapiens clone 24818 mRNA sequence.	Hs.106823
	Hs.106826
	Hs.10712
	Hs.107149
hypothetical protein	Hs.10729
Tachykinin Receptor 1	Hs.1080
glycophorin A	Hs.108694
Histone H1x	Hs.109804
CD66d	Hs.11
interleukin 17	Hs.110040
	Hs.110131
major histocompatibility complex, class I, F	Hs.110309
REV1	Hs.110347
HCR	Hs.110746
VWF	Hs.110802
high affinity immunoglobulin epsilon receptor beta subunit	Hs.11090
interleukin 22 receptor	Hs.110915
	Hs.110978
Homo sapiens ubiquitin specific protease 6 (Tre-2 oncogene) (USP6), mRNA.	Hs.111065
	Hs.111128
MMP2	Hs.111301
major histocompatibility complex, class II, DN alpha	Hs.11135
LTBR	Hs.1116
ESTs, Weakly similar to A41285 interleukin enhancer-binding factor ILF-1 [H.sapiens]	Hs.111941
Homo sapiens STRIN protein (STRIN), mRNA.	Hs.112144
MSH5	Hs.112193
TCRg	Hs.112259
	Hs.11307
CMKRL2	Hs.113207

Table 2: Candidate genes, Database mining

ferritin L chain	Hs.111334
TGF-b	Hs.1103
Homo sapiens ras homolog gene family, member H (ARHH), mRNA.	Hs.109918
lysosomal alpha-mannosidase (MANB)	Hs.108969
	Hs.108327
granzyme B	Hs.1051
HCC-4	Hs.10458
	Hs.10362
	Hs.102630
	Hs.101382
C4BPA	Hs.1012
CD125	Hs.100001
TERF2	Hs.100030
LIG3	Hs.100299
	Hs.157489
EST	Hs.157560
EST	Hs.157808
EST	Hs.157811
	Hs.158127
interleukin 18 receptor accessory protein	Hs.158315
CCR3	Hs.158324
Human DNA sequence from clone CTA-390C10 on chromosome 22q11.21-12.1 Contains an Immunoglobulin-like gene and a pseudogene similar to Beta Crystallin, ESTs, STSs, GSSs and taga and tat repeat polymorphisms	Hs.158352
ESTs	Hs.158576
	Hs.158874
EST	Hs.158875
EST	Hs.158876
EST	Hs.158878
EST	Hs.158956
EST	Hs.158967
EST	Hs.158969
EST	Hs.158971
EST	Hs.158988

CCR8	Hs.113222
LILRA3	Hs.113277
Human CXCR-5 (BLR-1)	Hs.113916
RAD51C	Hs.11393
myosin, heavy polypeptide 8, skeletal muscle, perinatal	Hs.113973
CD42a	Hs.1144
TNFRSF11A	Hs.114676
	Hs.114931
MSH4	Hs.115246
Homo sapiens dendritic cell immunoreceptor (DCIR), mRNA.	Hs.115515
REV3L (POLZ)	Hs.115521
JAK2	Hs.115541
OPG ligand	Hs.115770
PCDH12	Hs.115897
	Hs.166235
POLE1	Hs.166846
regulatory factor X, 5 (influences HLA class II expression)	Hs.166891
PIG-F (phosphatidyl-inositol-glycan class F)	Hs.166982
ESTs, Moderately similar to ILF1_HUMAN INTERLEUKIN ENHANCER-BINDING FACTOR 1 [H.sapiens]	Hs.167154
HLA-DRB6	Hs.167385
ret finger protein-like 3	Hs.167751
CD56	Hs.167988
RBT1	Hs.169138
APOE	Hs.169401
	Hs.16944
	Hs.169470
MMP12	Hs.1695
CD161	Hs.169824
tenascin XB	Hs.169886
	Hs.170027
	Hs.170150
C4A	Hs.170250

Table 2: Candidate genes, Database mining

CD120a=TNFR-1	Hs.159
EST	Hs.159000
	Hs.159013
EST	Hs.159025
EST	Hs.159059
IL18R1	Hs.159301
ftp-3	Hs.159494
CASP8	Hs.159651
EST	Hs.159655
EST	Hs.159660
EST	Hs.159678
kallikrein 12 (KLK12)	Hs.159679
EST	Hs.159682
EST	Hs.159683
EST	Hs.159693
EST	Hs.159706
EST	Hs.159718
SPO11	Hs.159737
EST	Hs.159754
EST	Hs.160401
EST	Hs.160405
EST	Hs.160408
EST	Hs.160410
EST	Hs.160423
RPA3	Hs.1608
ESTs	Hs.160946
EST	Hs.160956
ESTs	Hs.160978
EST	Hs.160980
EST	Hs.160981
EST	Hs.160982
EST	Hs.160983
Tachykinin Receptor 2	Hs.161305
RAD17 (RAD24)	Hs.16184
Human phosphatidylinositol 3-kinase catalytic subunit p110delta mRNA, complete cds.	Hs.162808
Human alpha-1 Ig germline C-region membrane-coding region, 3' end	Hs.163271
GCP-2	Hs.164021
	Hs.164284
EST	Hs.164331
	Hs.164427
	Hs.165568
ER	Hs.1657

TP53BP1	Hs.170263
ESTs	Hs.170274
ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION WARNING ENTRY□ [H.sapiens]	Hs.170338
ESTs	Hs.170578
EST	Hs.170579
ESTs	Hs.170580
EST	Hs.170581
ESTs	Hs.170583
EST	Hs.170586
EST	Hs.170588
EST	Hs.170589
	Hs.170772
ESTs	Hs.170786
EST	Hs.170909
EST	Hs.170912
EST	Hs.170933
ESTs	Hs.171004
EST	Hs.171095
EST	Hs.171098
ESTs	Hs.171101
EST	Hs.171108
ESTs	Hs.171110
ESTs	Hs.171113
ESTs	Hs.171117
EST	Hs.171119
ESTs	Hs.171120
EST	Hs.171122
EST	Hs.171123
EST	Hs.171124
EST	Hs.171140
EST	Hs.171216
EST	Hs.171260
ESTs	Hs.171264
RIP	Hs.171545
ESTs, Weakly similar to immunoglobulin superfamily member [D.melanogaster]	Hs.171697
CD22	Hs.171763
	Hs.171776
sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3C	Hs.171921
interleukin 11	Hs.1721
CD11b	Hs.172631
EST, Highly similar to APS [H.sapiens]	Hs.172656
ALK1	Hs.172670

Table 2: Candidate genes, Database mining

EST, Highly similar to JM26 [H.sapiens]	Hs.165701
EST	Hs.165702
EST	Hs.165704
EST	Hs.165732
regulatory factor X, 3 (influences HLA class II expression)	Hs.166019
LIG4	Hs.166091
TNFSF18	Hs.248197
EST	Hs.248228
H.sapiens rearranged gene for kappa immunoglobulin subgroup V kappa IV	Hs.248756
caspase 1, apoptosis-related cysteine protease (interleukin 1, beta, convertase)	Hs.2490
EST	Hs.249031
TNFRSF10A	Hs.249190
immunoglobulin lambda variable 3-10	Hs.249208
Homo sapiens mRNA for single-chain antibody, complete cds	Hs.249245
EST	Hs.250473
ESTs	Hs.250591
ESTs	Hs.250605
	Hs.25063
Human DNA sequence from clone RP1-149A16 on chromosome 22 Contains an IGLC (Immunoglobulin Lambda Chain C) pseudogene, the RFPL3 gene for Ret finger protein-like 3, the RFPL3S gene for Ret finger protein-like 3 antisense, the gene for a novel Immunoglobulin Lambda Chain V family protein, the gene for a novel protein similar to mouse RGDS (RALGDS, RALGEF, Guanine Nucleotide Dissociation Stimulator A) and rabbit oncogene RSC, the gene for a novel protein (ortholog of worm F16A11.2 and bacterial and archea-bacterial predicted proteins), the gene for a novel protein similar to BPI (Bacterial Permeability-Increasing Protein) and rabbit LBP (Liposaccharide-Binding Protein) and the 5' part of a novel gene. Contains ESTs, STSs, GSSs and three putative CpG islands	Hs.250675
ACE	Hs.250711
TREX2	Hs.251398

	Hs.172674
CD123	Hs.172689
ESTs	Hs.172822
Collal	Hs.172928
	Hs.172998
	Hs.173081
myosin, heavy polypeptide 3, skeletal muscle, embryonic	Hs.173084
	Hs.173201
Mediterranean fever (MEFV)	Hs.173730
	Hs.173749
interleukin 1 receptor accessory protein	Hs.173880
EST, Weakly similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.174231
EST	Hs.174242
EST	Hs.174300
EST	Hs.174634
EST	Hs.174635
EST	Hs.174650
EST	Hs.174673
EST	Hs.174716
EST	Hs.174740
EST	Hs.174778

Table 2: Candidate genes, Database mining

Human DNA sequence from clone 1170K4 on chromosome 22q12.2-13.1. Contains three novel genes, one of which codes for a Trypsin family protein with class A LDL receptor domains, and the IL2RB gene for Interleukin 2 Receptor, Beta (IL-2 Receptor, CD122 antigen). Contains a putative CpG island, ESTs, and GSSs	Hs.251417
EST	Hs.251539
EST	Hs.251540
C3	Hs.251972
EST	Hs.252273
EST	Hs.252359
ESTs, Moderately similar to T2DT_HUMAN TRANSCRIPTION INITIATION FACTOR TFIID 105 KDA SUBUNIT [H.sapiens]	Hs.252867
EST, Moderately similar to RS2_HUMAN 40S RIBOSOMAL PROTEIN S2 [H.sapiens]	Hs.253150
EST	Hs.253151
EST	Hs.253154
EST	Hs.253165
EST	Hs.253166
EST	Hs.253167
EST	Hs.253168
EST	Hs.253169
interleukin 1 receptor, type II	Hs.25333
	Hs.25361
EST	Hs.253742
EST	Hs.253743
EST, Weakly similar to AF161429_1 HSPC311 [H.sapiens]	Hs.253744
EST	Hs.253747
EST	Hs.253748
EST	Hs.253753
EST, Moderately similar to ALU5_HUMAN ALU SUBFAMILY SC SEQUENCE CONTAMINATION WARNING ENTRY□ [H.sapiens]	Hs.254108
ESTs	Hs.254948
ESTs	Hs.255011
EST	Hs.255118
EST	Hs.255119
EST	Hs.255123
EST	Hs.255129
EST	Hs.255134
EST	Hs.255135
EST	Hs.255139
EST	Hs.174779
EST, Weakly similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.174780
(KIAA0033) for ORF, partial cds.	Hs.174905
	Hs.175270
EST	Hs.175281
EST	Hs.175300
EST	Hs.175336
EST	Hs.175388
	Hs.175437
EST, Weakly similar to salivary proline-rich protein precursor [H.sapiens]	Hs.175777
EST	Hs.175803
ESTs	Hs.176337
EST	Hs.176374
EST	Hs.176380
EST	Hs.176404
EST	Hs.176406
LCK	Hs.1765
LIG1	Hs.1770
EST	Hs.177012
PERB11 family member in MHC class I region	Hs.17704
EST	Hs.177146
EST	Hs.177209
	Hs.177376
	Hs.177461
CD99	Hs.177543
PMS2	Hs.177548
human calmodulin	Hs.177656
	Hs.177712
Homo sapiens immunoglobulin lambda gene locus DNA, clone:288A10	Hs.178665
	Hs.178743
EST	Hs.179008
EST	Hs.179070
EST	Hs.179130

Table 2: Candidate genes, Database mining

EST	Hs.255140
ESTs	Hs.255142
EST	Hs.255150
EST	Hs.255152
ESTs	Hs.255153
ESTs	Hs.255157
ESTs	Hs.255171
EST	Hs.255172
EST, Moderately similar to PGTA_HUMAN RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT [H.sapiens]	Hs.255174
EST	Hs.255177
EST	Hs.255178
EST	Hs.255245
EST	Hs.255246
EST	Hs.255249
EST	Hs.255251
EST	Hs.255253
EST	Hs.255254
EST	Hs.255255
ESTs	Hs.255256
EST	Hs.255330
EST, Weakly similar to putative G protein-coupled Receptor [H.sapiens]	Hs.255333
EST	Hs.255336
EST	Hs.255337
EST	Hs.255339
EST	Hs.255340
EST	Hs.255341
ESTs	Hs.255343
EST	Hs.255347
EST	Hs.255349
EST	Hs.255350
EST	Hs.255354
ESTs	Hs.255359
ESTs	Hs.255387
EST	Hs.255388
EST	Hs.255389
ESTs	Hs.255390
EST	Hs.255392
EST	Hs.255444
EST	Hs.255446
EST	Hs.255448
ESTs	Hs.255449
EST	Hs.255454
EST	Hs.255455
EST	Hs.255457
EST	Hs.255459

EST	Hs.179132
	Hs.179149
EST	Hs.179490
EST	Hs.179492
promyelocytic leukemia cell mRNA, clones pHH58 and pHH81.	Hs.179735
	Hs.179817
major histocompatibility complex, class II, DO beta	Hs.1802
HLA-DRB1	Hs.180255
TNFRSF12	Hs.180338
RAD23A (HR23A)	Hs.180455
MKK3	Hs.180533
EST	Hs.180637
CD27	Hs.180841
STAT6	Hs.181015
TNFSF4	Hs.181097
immunoglobulin lambda locus	Hs.181125
	Hs.181368
CD3	Hs.181392
EST	Hs.255745
EST	Hs.255746
EST	Hs.255747
EST	Hs.255749
EST	Hs.255754
ESTs, Moderately similar to KIAA1271 protein [H.sapiens]	Hs.255759
EST	Hs.255762
EST	Hs.255763
EST	Hs.255764
EST	Hs.255766
EST	Hs.255767
EST	Hs.255768
EST	Hs.255769
EST	Hs.255770
EST	Hs.255772
EST	Hs.255777
EST	Hs.255778
EST	Hs.255779
EST	Hs.255782
EST	Hs.255783
EST	Hs.255784
EST	Hs.255785
EST, Weakly similar to Con1 [H.sapiens]	Hs.255788
EST	Hs.255791
EST	Hs.255794
EST	Hs.255796
EST	Hs.255797

Table 2: Candidate genes, Database mining

EST	Hs.255462
EST	Hs.255464
EST	Hs.255492
EST	Hs.255494
EST	Hs.255495
EST	Hs.255497
EST	Hs.255498
EST	Hs.255499
EST	Hs.255501
EST	Hs.255502
EST	Hs.255505
EST	Hs.255541
EST	Hs.255543
ESTs	Hs.255544
EST	Hs.255546
EST	Hs.255549
EST	Hs.255552
EST	Hs.255554
EST	Hs.255556
EST	Hs.255558
EST	Hs.255559
EST	Hs.255560
EST	Hs.255561
EST	Hs.255569
EST	Hs.255572
EST	Hs.255573
EST	Hs.255575
EST	Hs.255577
EST	Hs.255578
EST	Hs.255579
EST	Hs.255580
EST	Hs.255590
EST	Hs.255591
EST	Hs.255598
TNFRSF17	Hs.2556
EST	Hs.255600
EST	Hs.255601
ESTs, Highly similar to KIAA1039 protein [H.sapiens]	Hs.255603
EST	Hs.255614
EST	Hs.255615
ESTs	Hs.255617
EST	Hs.255618
EST	Hs.255621
EST	Hs.255622
ESTs	Hs.255625
EST	Hs.255626
ESTs	Hs.255627
ESTs	Hs.255630
EST	Hs.255632
EST	Hs.255633
EST	Hs.255799
ESTs	Hs.255877
EST	Hs.255880
EST	Hs.255920
EST	Hs.255927
CD40	Hs.25648
interleukin enhancer binding factor 3, 90kD	Hs.256583
ESTs	Hs.256810
EST	Hs.256956
EST	Hs.256957
EST	Hs.256959
EST	Hs.256961
EST	Hs.256970
EST	Hs.256971
ESTs	Hs.256979
ESTs	Hs.257572
EST	Hs.257579
EST	Hs.257581
EST	Hs.257582
EST	Hs.257630
EST	Hs.257632
EST	Hs.257633
EST	Hs.257636
EST	Hs.257640
ESTs	Hs.257641
EST	Hs.257644
EST	Hs.257645
EST	Hs.257646
EST	Hs.257647
EST	Hs.257667
EST	Hs.257668
EST	Hs.257677
EST	Hs.257679
EST	Hs.257680
ESTs	Hs.257682
ESTs	Hs.257684
EST	Hs.257687
EST	Hs.257688
EST	Hs.257690
EST	Hs.257695
EST	Hs.257697
EST	Hs.257705
EST	Hs.257706
EST	Hs.257709
ESTs, Moderately similar to ALU8_HUMAN ALU SUBFAMILY SX SEQUENCE CONTAMINATION WARNING ENTRY□ [H.sapiens]	Hs.257711
EST	Hs.257713
EST	Hs.257716
EST	Hs.257719
EST	Hs.257720
EST	Hs.257727

Table 2: Candidate genes, Database mining

EST	Hs.255634
EST	Hs.255635
EST	Hs.255637
ESTs	Hs.255639
EST	Hs.255641
EST	Hs.255644
EST	Hs.255645
EST	Hs.255646
EST	Hs.255647
EST	Hs.255648
EST	Hs.255649
EST	Hs.255650
EST	Hs.255653
EST	Hs.255657
EST	Hs.255661
ESTs	Hs.255664
EST	Hs.255665
EST	Hs.255666
EST	Hs.255668
EST	Hs.255671
EST	Hs.255672
EST	Hs.255673
EST	Hs.255674
EST	Hs.255675
EST	Hs.255677
EST	Hs.255679
EST	Hs.255681
EST	Hs.255682
EST	Hs.255686
ESTs	Hs.255687
EST	Hs.255688
ESTs	Hs.255689
EST	Hs.255691
EST	Hs.255692
ESTs	Hs.255693

EST	Hs.257730
EST	Hs.257738
EST	Hs.257743
ESTs	Hs.258513
EST	Hs.258820
EST	Hs.258864
sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4F	Hs.25887
EST	Hs.258898
EST	Hs.258933
interleukin 13 receptor, alpha 2	Hs.25954
Homo sapiens HSPC101 mRNA, partial cds	Hs.259683
EST	Hs.263695
ESTs	Hs.263784
TNFSF12	Hs.26401
EST	Hs.264154
EST	Hs.264654
CDw116b	Hs.265262
MHC binding factor, beta	Hs.2654
EST	Hs.265634
EST	Hs.266387
ESTs	Hs.268027
ATHS (LDLR?)	Hs.268571
ESTs, Highly similar to AAD18086 BAT2 [H.sapiens]	Hs.270193
ESTs	Hs.270198
ESTs	Hs.270294
ESTs, Weakly similar to alternatively spliced product using exon 13A [H.sapiens]	Hs.270542
ESTs, Moderately similar to ALU2_HUMAN ALU SUBFAMILY SB SEQUENCE CONTAMINATION WARNING ENTRY□ [H.sapiens]	Hs.270561
ESTs, Weakly similar to pro alpha 1(I) collagen [H.sapiens]	Hs.270564
ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION WARNING ENTRY□ [H.sapiens]	Hs.270578
ESTs, Moderately similar to brain-derived immunoglobulin superfamily molecule [M.musculus]	Hs.270588
TALL1	Hs.270737
ESTs	Hs.271206
MYH	Hs.271353
POLI (RAD30B)	Hs.271699
ADPRTL3	Hs.271742

Table 2: Candidate genes, Database mining

EST	Hs.255695	ESTs, Moderately similar to ALU8_HUMAN ALU SUBFAMILY SX SEQUENCE CONTAMINATION WARNING ENTRY□ [H.sapiens]	Hs.272075
EST, Highly similar to transmembrane chloride conductor protein [H.sapiens]	Hs.255697	Human DNA sequence from clone RP5-1170K4 on chromosome 22q12.2-13.1 Contains three novel genes, one of which codes for a Trypsin family protein with class A LDL receptor domains, and the IL2RB gene for Interleukin 2 Receptor, Beta (IL-2 Receptor, CD122 antigen), a	Hs.272271
EST	Hs.255698	interleukin 1 receptor accessory protein-like 2	Hs.272354
EST	Hs.255699	Homo sapiens partial IGVH3 V3-20 gene for immunoglobulin heavy chain V region, case 1, clone 2	Hs.272355
EST	Hs.255705	Homo sapiens partial IGVH3 gene for immunoglobulin heavy chain V region, case 1, clone 16	Hs.272356
EST	Hs.255706	Homo sapiens partial IGVH3 gene for immunoglobulin heavy chain V region, case 1, clone 19	Hs.272357
EST	Hs.255708	Homo sapiens partial IGVH3 gene for immunoglobulin heavy chain V region, case 1, cell Mo IV 72	Hs.272358
EST	Hs.255710	Homo sapiens partial IGVH1 gene for immunoglobulin heavy chain V region, case 1, cell Mo V 94	Hs.272359
EST	Hs.255713	Homo sapiens partial IGVL2 gene for immunoglobulin lambda light chain V region, case 1, cell Mo V 94	Hs.272360
EST	Hs.255717	Homo sapiens partial IGVH3 gene for immunoglobulin heavy chain V region, case 1, cell Mo VI 7	Hs.272361
EST	Hs.255718	Homo sapiens partial IGVL1 gene for immunoglobulin lambda light chain V region, case 1, cell Mo VI 65	Hs.272362
EST	Hs.255721	Homo sapiens partial IGVH3 gene for immunoglobulin heavy chain V region, case 1, cell Mo VI 162	Hs.272363
ESTs	Hs.255723	Homo sapiens partial IGVH3 DP29 gene for immunoglobulin heavy chain V region, case 1, cell Mo VII 116	Hs.272364
EST	Hs.255725	Homo sapiens partial IGVH4 gene for immunoglobulin heavy chain V region, case 2, cell D 56	Hs.272365
EST	Hs.255726	Homo sapiens partial IGVH3 gene for immunoglobulin heavy chain V region, case 2, cell E 172	Hs.272366
EST	Hs.255727	interleukin 20	Hs.272373

Table 2: Candidate genes, Database mining

EST	Hs.255736
EST	Hs.255740
EST	Hs.255742
EST	Hs.255743
EST	Hs.7569
SMAD4	Hs.75862
Homo sapiens splicing factor, arginine/serine-rich 4 (SFRS4) mRNA.	Hs.76122
thymosin beta-10	Hs.76293
CD63	Hs.76294
AIF1	Hs.76364
phospholipase A2, group IIA (platelets, synovial fluid),	Hs.76422
CES1	Hs.76688
ubiquitin conjugating enzyme	Hs.76932
Homo sapiens KIAA0963 protein (KIAA0963), mRNA.	Hs.7724
Homo sapiens fragile histidine triad gene (FHIT) mRNA.	Hs.77252
PAF-AH	Hs.77318
Mig	Hs.77367
DDB2	Hs.77602
ATR	Hs.77613
XPB (ERCC3)	Hs.77929
PNKP	Hs.78016
C7	Hs.78065

Human DNA sequence from clone RP1-149A16 on chromosome 22 Contains an IGLC (Immunoglobulin Lambda Chain C) pseudogene, the RFPL3 gene for Ret finger protein-like 3, the RFPL3S gene for Ret finger protein-like 3 antisense, the gene for a novel Immunoglobulin Lambda Chain V family protein, the gene for a novel protein similar to mouse RGDS (RALGDS, RALGEF, Guanine Nucleotide Dissociation Stimulator A) and rabbit oncogene RSC, the gene for a novel protein (ortholog of worm F16A11.2 and bacterial and archea-bacterial predicted proteins), the gene for a novel protein similar to BPI (Bacterial Permeability-Increasing Protein) and rabbit LBP (Liposaccharide-Binding Protein) and the 5' part of a novel gene. Contains ESTs, STSs, GSSs and three putative CpG islands	Hs.272521
TdT	Hs.272537
ret finger protein-like 3 antisense	Hs.274285
PRKR	Hs.274382
H.sapiens immunoglobulin epsilon chain	Hs.274600
EST, Weakly similar to HLA-DQ alpha chain [H.sapiens]	Hs.275720
EST, Weakly similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.276279
EST	Hs.276341
EST	Hs.276342
EST, Weakly similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.276353
EST	Hs.276774
EST	Hs.276819
EST	Hs.276871
EST, Weakly similar to FBRL_HUMAN FIBRILLARIN [H.sapiens]	Hs.276872
EST	Hs.276887
EST	Hs.276902
EST	Hs.276917
EST	Hs.276918
EST, Weakly similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.276938
EST	Hs.277051
EST	Hs.277052
EST, Moderately similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.277236

Table 2: Candidate genes, Database mining

Homo sapiens small nuclear RNA activating complex, polypeptide 2, 45kD (SNAPC2) mRNA.	Hs.78403
	Hs.78465
sphingolipid activator protein / cerebroside sulfate activator protein	Hs.78575
Homo sapiens aminolevulinate, delta-, synthase 1 (ALAS1), nuclear gene encoding mitochondrial protein, mRNA.	Hs.78712
tyrosine kinase with immunoglobulin and epidermal growth factor homology domains	Hs.78824
Hsp72	Hs.78846
UNG	Hs.78853
CX3CR1	Hs.78913
MSH2	Hs.78934
CRHR1	Hs.79117
BCL2	Hs.79241
P-selectin	Hs.79283
UBE2VE (MMS2)	Hs.79300
retinoid X receptor, beta	Hs.79372
MPG	Hs.79396
RPA2	Hs.79411
heat shock 70kD protein-like 1	Hs.80288
FANCG (XRCC9)	Hs.8047
CD43	Hs.80738
POLG	Hs.80961
Human CB-4 transcript of unrearranged immunoglobulin V(H)5 gene	Hs.81220
Human L2-9 transcript of unrearranged immunoglobulin V(H)5 pseudogene	Hs.81221
immunoglobulin superfamily, member 3	Hs.81234
UBL1	Hs.81424
PF4	Hs.81564
palmitoyl-protein thioesterase 2	Hs.81737
natural killer cell receptor, immunoglobulin superfamily member	Hs.81743
TNFRSF11B	Hs.81791
interleukin 6 signal transducer (gp130, oncostatin M receptor)	Hs.82065
CD138	Hs.82109
Human monocytic leukaemia zinc finger protein (MOZ) mRNA, complete cds.	Hs.82210
sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B	Hs.82222
HPRT	Hs.82314
Human RNA binding protein Etr-3 mRNA, complete cds.	Hs.82321
MNAT1	Hs.82380

EST, Moderately similar to DEAD Box Protein 5 [H.sapiens]	Hs.277237
EST	Hs.277238
EST	Hs.277286
major histocompatibility complex, class I, C	Hs.277477
EST, Weakly similar to AF150959 1 immunoglobulin G1 Fc fragment [H.sapiens]	Hs.277591
EST	Hs.277714
EST	Hs.277715
EST	Hs.277716
EST	Hs.277717
EST	Hs.277718
EST, Weakly similar to BAT3_HUMAN LARGE PROLINE-RICH PROTEIN BAT3 [H.sapiens]	Hs.277774
EST	Hs.277975
EST	Hs.278060
cytochrome P450, subfamily XXIA (steroid 21-hydroxylase, congenital adrenal hyperplasia), polypeptide 2	Hs.278430
KIAA0015 gene product	Hs.278441
CD32B	Hs.278443
KIR2DL1	Hs.278453
CD158a	Hs.278455
CD24	Hs.278667
HLA class II region expressed gene KE4	Hs.278721
IL-17C	Hs.278911
HSPC048 protein (HSPC048)	Hs.278944
HSPC054 protein (HSPC054)	Hs.278946
HSPC073 protein (HSPC073)	Hs.278948
ESTs	Hs.279066
ESTs	Hs.279067
ESTs	Hs.279068
ESTs	Hs.279069
ESTs	Hs.279070
ESTs	Hs.279071
ESTs	Hs.279072
ESTs, Weakly similar to KIAA0052 protein [H.sapiens]	Hs.279073
ESTs	Hs.279074
ESTs	Hs.279075
ESTs	Hs.279076

Table 2: Candidate genes, Database mining

SMAD2	Hs.82483
CD47	Hs.82685
CETN2	Hs.82794
protein phosphatase 1, regulatory (inhibitor) subunit 11	Hs.82887
MMP1	Hs.83169
D3-type cyclin (CCND3)	Hs.83173
MMP3	Hs.83326
TNFSF10	Hs.83429
CD33	Hs.83731
CD102	Hs.83733
	Hs.84153
interleukin 8 receptor, beta	Hs.846
titin immunoglobulin domain protein (myotilin)	Hs.84665
KU80 (XRCC5)	Hs.84981
Raf-1	Hs.85181
major histocompatibility complex, class I, J (pseudogene)	Hs.85242
RELB	Hs.858
	Hs.85923
ERK1	Hs.861
FADD	Hs.86131
MHC class I polypeptide-related sequence A	Hs.90598
TNF receptor-associated factor 6	Hs.90957
Topo3A	Hs.91175
PARG	Hs.91390
HLA-DPA1	Hs.914
SEEK1	Hs.91600
POLD1	Hs.99890
ALK4	Hs.99954
XPD (ERCC2)	Hs.99987
SCYA25 (CCL25)	Hs.50404
SCYA19 (CCL19)	Hs.50002
TCIRG1	Hs.46465
PAF-Receptor	Hs.46
CD26	Hs.44926
	Hs.44865
REL	Hs.44313
IL-17	Hs.41724
CD49d	Hs.40034
CCR2	Hs.395
	Hs.3688
TNF-b	Hs.36
lactoferrin	Hs.347
MCP-1	Hs.340
CD150	Hs.32970
IL-10Ra	Hs.327
EGR1	Hs.326035

ESTs	Hs.279077
EST	Hs.279078
EST	Hs.279079
ESTs	Hs.279080
EST	Hs.279081
ESTs	Hs.279082
ESTs	Hs.279083
ESTs	Hs.279084
ESTs	Hs.279085
ESTs	Hs.279086
ESTs, Weakly similar to AF201422_1 splicing coactivator subunit SRm300 [H.sapiens]	Hs.279087
ESTs	Hs.279088
ESTs	Hs.279089
	Hs.86437
	Hs.86761
CD118 = IFNAR-2	Hs.86958
	Hs.87113
PGHS-1	Hs.88474
	Hs.8882
LT-b	Hs.890
EST	Hs.92440
	Hs.92460
myosin-binding protein H	Hs.927
IFN-b	Hs.93177
C8A	Hs.93210
pre-B-cell leukemia transcription factor 2	Hs.93728
Tachykinin Receptor 3	Hs.942
Homo sapiens cDNA FLJ12242 fis, clone MAMMA1001292	Hs.94810
CD29	Hs287797
LIF	Hs.2250
Human IP-10	Hs.2248
IL-5	Hs.2247
G-CSF	Hs.2233
TGF-bR	Hs.220
G-CSFR	Hs.2175
CD15	Hs.2173
STAT1	Hs.21486
CD85	Hs.204040
HCC-1	Hs.20144
Fas ligand	Hs.2007
CD28	Hs.1987
HLA-DQA1	Hs.198253
Ku70 (G22P1)	Hs.197345
PGHS-2	Hs.196384
CDw128	Hs.194778
IL-10	Hs.193717

Table 2: Candidate genes, Database mining

SCYC1 (XCL1)	Hs.3195
HLA-DR	Hs.318720
Topo I (TOP1)	Hs.317
SCYA2 (MCP1)	Hs.303649
HuRNPD	Hs.303627
Human C mu gene for IgM heavy chain exons CH1-4, secretory	Hs.302063
P1	Hs.297681
immunoglobulin lambda joining 3	Hs.289110
major histocompatibility complex, class II, DQ alpha 2	Hs.289095
HSPCA	Hs.289088
interleukin 22	Hs.287369
ribosomal protein L4	Hs.286
IgM	Hs.285823
EST	Hs.283267
TREM1	Hs.283022
HLA-DRB3	Hs.279930
LIFR	Hs.2798
C4B	Hs.278625
EST	Hs.276907
CDw52	Hs.276770
CD16 b	Hs.274467
heat shock 70kD protein 1B	Hs.274402
Th1	Hs.273385
MIP-5/HCC-2	Hs.272493
TBX21	Hs.272409
Homo sapiens mRNA; cDNA DKFZp434O2417 (from clone DKFZp434O2417); partial cds	Hs.272307
Human DNA sequence from clone RP1-108C2 on chromosome 6p12.1-21.1. Contains the MCM3 gene for minichromosome maintenance deficient (S. cerevisiae) 3 (DNA replication licensing factor, DNA polymerase alpha holoenzyme-associated protein P1, RLF beta subunit), a CACT (carnitine/acylcarnitine translocase) pseudogene, part of the gene for a PUTATIVE novel protein similar to IL17 (interleukin 17 (cytotoxic T-lymphocyte-associated serine esterase 8)) (cytotoxic T lymphocyte-associated antigen 8, CTLA8), ESTs, STSs, GSSs and a putative CpG island	Hs.272295

CD126	Hs.193400
	Hs.1880
CD98	Hs.184601
	Hs.184542
MHC class I region ORF	Hs.1845
CDw116a	Hs.182378
HLA-DRB5	Hs.181366
major histocompatibility complex, class I, A	Hs.181244
elongation factor 1-alpha (clone CEF4)	Hs.181165
CD119	Hs.180866
	Hs.180804
	Hs.180532
POLB	Hs.180107
CD1d	Hs.1799
CD87	Hs.179657
minichromosome maintenance deficient (S. cerevisiae) 3	Hs.179565
RAD23B (HR23B)	Hs.178658
	Hs.178391
	Hs.177781
ADPRT	Hs.177766
IFNGR2	Hs.177559
CD16 a	Hs.176663
CD4	Hs.17483
SCYC2 (XCL2)	Hs.174228
CD115	Hs.174142
CD11a	Hs.174103
IL-10Rb	Hs.173936

Table 2: Candidate genes, Database mining

CD49b	Hs.271986
MCP-2	Hs.271387
CD49c	Hs.265829
NBS1	Hs.25812
CD120b = TNFRSF1B	Hs.256278
CDw75	Hs.2554
CD82	Hs.25409
MCP-3	Hs.251526
xanthine oxidase	Hs.250
Human Ig rearranged lambda-chain mRNA, subgroup VL3, V-J region, partial cds	Hs.247947
Eotaxin-2/MPIF-2	Hs.247838
CTLA-4	Hs.247824
immunoglobulin kappa variable 1-9	Hs.247792
CD68	Hs.246381
OSMR	Hs.238648
CDw127	Hs.237868
transcription factor 8 (represses interleukin 2 expression)	Hs.232068
CD8b	Hs.2299
EST	Hs.229374
TRF4-1	Hs.225951
CD3g	Hs.2259
C2	Hs.2253
	Hs.116834
	Hs.117741
Human MHC Class I region proline rich protein mRNA, complete cds	Hs.118354
ESTs, Weakly similar to FCE2 MOUSE LOW AFFINITY IMMUNOGLOBULIN EPSILON FC RECEPTOR [M.musculus]	Hs.118392
MKK6	Hs.118825
	Hs.118895
H.sapiens mRNA for ITBA4 gene.	Hs.119018
	Hs.119057
TNFRSF10c	Hs.119684
	Hs.12064
	Hs.120907
acid phosphatase 5, tartrate resistant	Hs.1211
	Hs.121297

MSCF	Hs.173894
TDG	Hs.173824
RAC1	Hs.173737
integrin cytoplasmic domain-associated protein 1	Hs.173274
IL2R	Hs.1724
IL-1a	Hs.1722
	Hs.171872
	Hs.171118
EST	Hs.171009
EST	Hs.170934
EST	Hs.170587
IL-9R	Hs.1702
CD45	Hs.170121
TGF-a	Hs.170009
CD44	Hs.169610
Fyn	Hs.169370
MPIF-1	Hs.169191
ICAM-1	Hs.168383
IL-15	Hs.168132
STAT5A	Hs.167503
ESTs	Hs.167208
ESTs	Hs.165693
	Hs.135750
DINB1 (POLK)	Hs.135756
Human DNA sequence from clone RP1-238O23 on chromosome 6. Contains part of the gene for a novel protein similar to PIGR (polymeric immunoglobulin receptor), part of the gene for a novel protein similar to rat SAC (soluble adenylyl cyclase), ESTs, STSs and GSS	Hs.136141
	Hs.136254
	Hs.13646
	Hs.136537
Histone H1 (F3)	Hs.136857
MGMT	Hs.1384
	Hs.138563
IgG	Hs.140
	Hs.140478
	Hs.14070
	Hs.141153

Table 2: Candidate genes, Database mining

Human immunoglobulin (mAb59) light chain V region mRNA, partial sequence	Hs.121508
IL12Rb1	Hs.121544
Human MHC class II DO-alpha mRNA, partial cds	Hs.123041
Histone H4 (H4F2)	Hs.123053
TSHR	Hs.123078
	Hs.123445
regulatory factor X, 1 (influences HLA class II expression)	Hs.123638
CD13	Hs.1239
IL-15R	Hs.12503
RAD51L3 (RAD51D)	Hs.125244
CDw90	Hs.125359
LYPLA1	Hs.12540
ESTs, Weakly similar to AF201951 1 high affinity immunoglobulin epsilon receptor beta subunit [H.sapiens]	Hs.126580
	Hs.127128
	Hs.127444
C5	Hs.1281
C8G	Hs.1285
RAD54B	Hs.128501
	Hs.129020
	Hs.129268
	Hs.129332
XRCC2	Hs.129727
potassium voltage-gated channel, Shaw-related subfamily, member 3 (KCNC3)	Hs.129738
interleukin 17 receptor	Hs.129751
CD134	Hs.129780
TNFRSF10d	Hs.129844
POLL	Hs.129903
GADD153=growth arrest and DNA-damage-inducible gene / fus-chop fusion protein	Hs.129913

	Hs.143954
ESTs, Moderately similar to I1BC_HUMAN INTERLEUKIN-1 BETA CONVERTASE PRECURSOR [H.sapiens]	Hs.144814
CHK2 (Rad53)	Hs.146329
EST	Hs.146591
	Hs.147040
CD42b	Hs.1472
	Hs.149235
AICD	Hs.149342
Homo sapiens putative tumor suppressor protein (101F6) mRNA, complete cds.	Hs.149443
CD49e	Hs.149609
heparan sulfate proteoglycan (HSPG) core protein	Hs.1501
CD107a	Hs.150101
ESTs, Weakly similar to I57587 MHC HLA-SX-alpha [H.sapiens]	Hs.150175
ALK2	Hs.150402
WRN	Hs.150477
EST	Hs.150708
XRCC4	Hs.150930
IFN-a	Hs.1510
MAPK	Hs.151051
	Hs.15200
immunoglobulin mu binding protein 2	Hs.1521
4-1BBL	Hs.1524
	Hs.152818
HUS1	Hs.152983
SWAP70	Hs.153026
DOM-3 (C. elegans) homolog Z	Hs.153299
	Hs.153551
	Hs.15370

Table 2: Candidate genes, Database mining

solute carrier family 5 (neutral amino acid transporters, system A), member 4	Hs.130101
	Hs.130232
	Hs.13034
CD30L	Hs.1313
SCYA26 (CCL26)	Hs.131342
CD30	Hs.1314
	Hs.131885
	Hs.131887
	Hs.13256
ESTs	Hs.132775
Homo sapiens (clone 3.8-1) MHC class I mRNA fragment	Hs.132807
	Hs.13288
	Hs.132943
EST	Hs.133261
	Hs.133388
EST	Hs.133393
EST	Hs.133930
ESTs	Hs.133947
ESTs	Hs.133949
EST	Hs.134017
EST	Hs.134018
EST	Hs.134590
	Hs.135135
immunoglobulin superfamily, member 6	Hs.135194
	Hs.135570
Homo sapiens arrestin, beta 2 (ARRB2) mRNA.	Hs.18142
myeloperoxidase	Hs.1817
APO-1	Hs.182359
TRAP1	Hs.182366
	Hs.182594
TNFRSF16	Hs.1827
	Hs.182817
regulatory factor X, 4 (influences HLA class II expression)	Hs.183009
Homo sapiens killer cell lectin-like receptor F1 (KLRF1), mRNA.	Hs.183125
	Hs.183171
EST	Hs.183386
	Hs.183656
	Hs.18368

SMAD6	Hs.153863
APEXL2	Hs.154149
	Hs.154198
	Hs.154366
BCL6	Hs.155024
	Hs.155150
	Hs.155402
RAIDD	Hs.155566
POLH	Hs.155573
	Hs.15589
Homo sapiens mRNA for KIAA0695 protein, complete cds.	Hs.155976
SNM1 (PS02)	Hs.1560
Topo2A	Hs.156346
ESTs, Highly similar to MHC class II antigen [H.sapiens]	Hs.156811
Histamine H1 receptor	Hs.1570
	Hs.157118
	Hs.157267
EST	Hs.157279
EST	Hs.157280
EST	Hs.157308
EST	Hs.157309
EST	Hs.157310
EST	Hs.157311
ESTs	Hs.157344
ret finger protein-like 2	Hs.157427
	Hs.214956
WASP	Hs.2157
CD88	Hs.2161
	Hs.21618
ring finger protein 5	Hs.216354
class II cytokine receptor ZCYTOR7	Hs.21814
	Hs.219149
cyclophilin-related protein	Hs.219153
Homo sapiens mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase (MGAT2) mRNA.	Hs.219479
perforin	Hs.2200
	Hs.220154
ESTs, Weakly similar to FCE2 MOUSE LOW AFFINITY IMMUNOGLOBULIN EPSILON FC RECEPTOR [M.musculus]	Hs.220649
	Hs.220868

Table 2: Candidate genes, Database mining

advanced glycosylation end product-specific receptor	Hs.184
CDK7	Hs.184298
	Hs.184376
CCR4	Hs.184926
EST, Weakly similar to A27307 proline-rich phosphoprotein [H.sapiens]	Hs.185463
EST	Hs.185498
EST, Weakly similar to B39066 proline-rich protein 15 - rat [R.norvegicus]	Hs.186243
EST, Weakly similar to salivary proline-rich protein [R.norvegicus]	Hs.186265
EST	Hs.187200
	Hs.188048
EST	Hs.188075
EST	Hs.188194
EST	Hs.188300
	Hs.190251
	Hs.19056
EST	Hs.190831
MAPK8	Hs.190913
EST	Hs.190921
EST, Weakly similar to S39206 hypothetical protein 1 - rat□ [R.norvegicus]	Hs.190924
GTF2H2	Hs.191356
	Hs.191367
	Hs.191914
ESTs, Weakly similar to immunoglobulin superfamily member [D.melanogaster]	Hs.192078
XPA	Hs.192803
CD89	Hs.193122
DFFRY	Hs.193145
CD35	Hs.193716
REV7 (MAD2L2)	Hs.19400
	Hs.194082
	Hs.194110
BRCA1	Hs.194143
ESTs, Moderately similar to MHC Class I region proline rich protein [H.sapiens]	Hs.194249
	Hs.194534
Topo3B	Hs.194685
	Hs.220960
immunoglobulin superfamily, member 1	Hs.22111
	Hs.221539
ESTs	Hs.221694
	Hs.222921
	Hs.222942
EST	Hs.223520
EST	Hs.223935
EST, Moderately similar to SMO_HUMAN SMOOTHENED HOMOLOG PRECURSOR [H.sapiens]	Hs.224178
Blk	Hs.2243
EST	Hs.224344
EST	Hs.224408
EST	Hs.224409
CPN1	Hs.2246
MMP7	Hs.2256
MMP10	Hs.2258
CCR9	Hs.225946
toll-like receptor 6 (TLR6)	Hs.227105
XPR1	Hs.227656
CD49f	Hs.227730
	Hs.22790
EST	Hs.228337
EST, Highly similar to 1409218A elastase [H.sapiens]	Hs.228525
EST	Hs.228528
EST, Moderately similar to R37A_HUMAN 60S RIBOSOMAL PROTEIN L37A□ [H.sapiens]	Hs.228874
EST	Hs.228891
EST	Hs.228926
EST	Hs.229071
EST	Hs.229405
EST	Hs.229494
EST, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION WARNING ENTRY□ [H.sapiens]	Hs.229560
EST, Moderately similar to AAD18086 BAT2 [H.sapiens]	Hs.229901
EST	Hs.229902
EST, Highly similar to 1409218A elastase [H.sapiens]	Hs.230053

Table 2: Candidate genes, Database mining

Human DNA sequence from clone 1170K4 on chromosome 22q12.2-13.1. Contains three novel genes, one of which codes for a Trypsin family protein with class A LDL receptor domains, and the IL2RB gene for Interleukin 2 Receptor, Beta (IL-2 Receptor, CD122 antigen). Contains a putative CpG island, ESTs, and GSSs	Hs.194750	RAD51	Hs.23044
major histocompatibility complex, class II, DP alpha 2 (pseudogene)	Hs.194764	EST, Moderately similar to A54746 adhalin precursor - human [H.sapiens]	Hs.230485
Human DNA sequence from clone RP11-367J7 on chromosome 1. Contains (part of) two or more genes for novel Immunoglobulin domains containing proteins, a SON DNA binding protein (SON) pseudogene, a voltage-dependent anion channel 1 (VDAC1) (plasmalemmal porin) pseudogene, ESTs, STSs and GSSs	Hs.194976	EST	Hs.230691
	Hs.195447	EST	Hs.230775
PDGF-B	Hs.1976	EST	Hs.230805
CXCR3	Hs.198252	EST	Hs.230848
	Hs.198694	EST	Hs.230862
	Hs.198738	EST	Hs.230874
MAR/SAR DNA binding protein (SATB1)	Hs.198822	EST	Hs.230931
CHUK	Hs.198998	EST	Hs.231031
hemochromatosis	Hs.20019	EST	Hs.231261
T-cell receptor active beta-chain	Hs.2003	EST	Hs.231284
APO-1	Hs.2007,	EST	Hs.231285
RXRA	Hs.20084	EST	Hs.231292
EST	Hs.200876	EST, Weakly similar to putative mitochondrial outer membrane protein import receptor [H.sapiens]	Hs.231512
	Hs.201194	Homo sapiens mRNA for KIAA0529 protein, partial cds.	Hs.23168
TCRd	Hs.2014	EST	Hs.235042
ESTs, Highly similar to TNF-alpha converting enzyme [H.sapiens]	Hs.202407	EST	Hs.235826
	Hs.202608	TREX1 (Dnase III)	Hs.23595
Integrin b1 = CD29	Hs.202661	EST	Hs.237126
thrombomodulin	Hs.2030		Hs.23860
	Hs.203064	RAD9	Hs.240457
	Hs.203184	1-acylglycerol-3-phosphate O-acyltransferase 1 (lysophosphatidic acid acyltransferase, alpha)	Hs.240534
	Hs.203584	EST	Hs.240635
EST	Hs.204477	EST, Weakly similar to ALU8_HUMAN ALU SUBFAMILY SX SEQUENCE CONTAMINATION WARNING ENTRY [H.sapiens]	Hs.241136

Table 2: Candidate genes, Database mining

EST	Hs.204480
EST, Weakly similar to CA13_HUMAN COLLAGEN ALPHA 1(III) CHAIN PRECURSOR□ [H.sapiens]	Hs.204483
ESTs	Hs.204588
EST, Weakly similar to salivary proline-rich protein 1 [H.sapiens]	Hs.204598
EST	Hs.204610
ESTs	Hs.204703
ESTs	Hs.204751
EST	Hs.204760
EST	Hs.204771
ESTs	Hs.204873
ESTs	Hs.204932
EST	Hs.204954
EST	Hs.205158
ESTs	Hs.205159
ESTs	Hs.205327
CD39	Hs.205353
ESTs	Hs.205435
EST	Hs.205438
EST, Highly similar to elastic titin [H.sapiens]	Hs.205452
EST	Hs.205456
MRE11A	Hs.20555
HLA class II region expressed gene KE2	Hs.205736
EST	Hs.205788
ESTs	Hs.205789
EST	Hs.205803
EST	Hs.205815
ESTs	Hs.206160
	Hs.206654
EST	Hs.207060
EST	Hs.207062
EST	Hs.207063
EST	Hs.207473
ESTs	Hs.207474

TNFSF15	Hs.241382
interleukin 1 receptor accessory protein-like 1	Hs.241385
RANTES	Hs.241392
sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	Hs.2414
POLQ	Hs.241517
TNF-a	Hs.241570
Homo sapiens genes encoding RNCC protein, DDAH protein, Ly6-C protein, Ly6-D protein and immunoglobulin receptor	Hs.241586
megakaryocyte-enhanced gene transcript 1 protein	Hs.241587
EST, Moderately similar to 1409218A elastase [H.sapiens]	Hs.241981
EST	Hs.241982
EST	Hs.241983
EST	Hs.242605
ADPRT2	Hs.24284
EST	Hs.243284
EST	Hs.243286
ESTs	Hs.243288
SCYB14	Hs.24395
EST	Hs.244046
EST	Hs.244048
EST	Hs.244049
EST	Hs.244050
RFXAP	Hs.24422
	Hs.24435
STAT5B	Hs.244613
EST	Hs.244666
EST	Hs.245586
CDw108	Hs.24640
ESTs	Hs.246796
dimethylarginine dimethylaminohydrolase 2	Hs.247362
Homo sapiens clone mcg53-54 immunoglobulin lambda light chain variable region 4a mRNA, partial cds	Hs.247721
Homo sapiens ELK1 pseudogene (ELK2) and immunoglobulin heavy chain gamma pseudogene (IGHGP)	Hs.247775
immunoglobulin kappa variable 1/OR2-108	Hs.247804
butyrophilin-like 2 (MHC class II associated)	Hs.247808

Table 2: Candidate genes, Database mining

ESTs	Hs.207971
EST	Hs.207993
EST	Hs.208153
EST, Weakly similar to S10889 proline-rich protein - human□ [H.sapiens]	Hs.208667
ESTs	Hs.209142
EST	Hs.209261
ESTs	Hs.209306
	Hs.209362
EST, Weakly similar to FCEB MOUSE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT [M.musculus]	Hs.209540
EST	Hs.209913
EST	Hs.209989
EST	Hs.210049
EST, Moderately similar to probable sodium potassium ATPase gamma chain [H.sapiens]	Hs.210276
EST, Weakly similar to N-WASP [H.sapiens]	Hs.210306
EST	Hs.210307
EST	Hs.210385
interleukin 21 receptor	Hs.210546
EST	Hs.210727

Homo sapiens genes encoding RNCC protein, DDAH protein, Ly6-C protein, Ly6-D protein and immunoglobulin receptor	Hs.247879
Histamine H2 receptor	Hs.247885
Human anti-streptococcal/anti-myosin immunoglobulin lambda light chain variable region mRNA, partial cds	Hs.247898
Homo sapiens isolate donor Z clone Z55K immunoglobulin kappa light chain variable region mRNA, partial cds	Hs.247907
Homo sapiens isolate donor D clone D103L immunoglobulin lambda light chain variable region mRNA, partial cds	Hs.247908
Homo sapiens isolate 459 immunoglobulin lambda light chain variable region (IGL) gene, partial cds	Hs.247909
Homo sapiens isolate donor N clone N88K immunoglobulin kappa light chain variable region mRNA, partial cds	Hs.247910
Homo sapiens isolate donor N clone N8K immunoglobulin kappa light chain variable region mRNA, partial cds	Hs.247911
Human Ig rearranged mu-chain V-region gene, subgroup VH-III, exon 1 and 2	Hs.247923
Epsilon , IgE=membrane-bound IgE, epsilon m/s isoform {alternative splicing} [human, mRNA Partial, 216 nt]	Hs.247930
H.sapiens (T1.1) mRNA for IG lambda light chain	Hs.247949
H.sapiens mRNA for Ig light chain, variable region (ID:CLL001VL)	Hs.247950
Human interleukin 2 gene, clone pATtacIL-2C/2TT, complete cds, clone pATtacIL-2C/2TT	Hs.247956
pre-B lymphocyte gene 1	Hs.247979
Human immunoglobulin heavy chain variable region (V4-31) gene, partial cds	Hs.247987
Human immunoglobulin heavy chain variable region (V4-30.2) gene, partial cds	Hs.247989
Human DNA sequence from phage LAW2 from a contig from the tip of the short arm of chromosome 16, spanning 2Mb of 16p13.3 Contains Interleukin 9 receptor pseudogene	Hs.247991
Homo sapiens HLA class III region containing NOTCH4 gene, partial sequence, homeobox PBX2 (HPBX) gene, receptor for advanced glycosylation end products (RAGE) gene, complete cds, and 6 unidentified cds	Hs.247993

Table 2: Candidate genes, Database mining

	Hs.211266
SMAD3	Hs.211578
MHC class I polypeptide-related sequence B	Hs.211580
ESTs, Weakly similar to CA1B_MOUSE COLLAGEN ALPHA 1(XI) CHAIN PRECURSOR□ [M.musculus]	Hs.211744
sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E	Hs.212414
TNFRSF18	Hs.212680
Homo sapiens general transcription factor 2-I pseudogene 1 (GTF2IP1) mRNA.	Hs.212939
RAD18	Hs.21320
	Hs.213226
ESTs	Hs.279090
ESTs	Hs.279091
ESTs	Hs.279092
EST	Hs.279093
ESTs	Hs.279094
ESTs	Hs.279095
ESTs, Weakly similar to AF279265_1 putative anion transporter 1 [H.sapiens]	Hs.279096
ESTs	Hs.279097
EST	Hs.279098
ESTs	Hs.279099
ESTs	Hs.279100
ESTs	Hs.279101
ESTs	Hs.279102
ESTs	Hs.279103
ESTs	Hs.279104
ESTs	Hs.279105
ESTs	Hs.279106
EST	Hs.279107
ESTs	Hs.279108
EST	Hs.279109
ESTs	Hs.279110

Homo sapiens immunoglobulin lambda gene locus DNA, clone:61D6	Hs.248010
immunoglobulin lambda variable 9-49	Hs.248011
immunoglobulin lambda variable 4-3	Hs.248012
H.sapiens mRNA for IgG lambda light chain V-J-C region (clone Tgl11)	Hs.248030
Human immunoglobulin (mAb56) light chain V region mRNA, partial sequence	Hs.248043
Homo sapiens lymphocyte-predominant Hodgkin's disease case #4 immunoglobulin heavy chain gene, variable region, partial cds	Hs.248077
Homo sapiens lymphocyte-predominant Hodgkin's disease case #7 immunoglobulin heavy chain gene, variable region, partial cds	Hs.248078
Homo sapiens clone ASMneg1-b3 immunoglobulin lambda chain VJ region, (IGL) mRNA, partial cds	Hs.248083
OSM	Hs.248156
	Hs.29128
Homo sapiens clone 24659 mRNA sequence.	Hs.29206
EST	Hs.292235
EST	Hs.292450
EST, Moderately similar to Ewing sarcoma breakpoint region 1, isoform EWS [H.sapiens]	Hs.292455
EST	Hs.292461
ESTs	Hs.292501
EST	Hs.292516
EST	Hs.292517
EST	Hs.292520
EST, Moderately similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.292540
EST	Hs.292545
EST, Weakly similar to ORFII [H.sapiens]	Hs.292704
EST	Hs.292761
ESTs	Hs.292803
ESTs	Hs.293183
ESTs	Hs.293280
ESTs	Hs.293281
ESTs, Moderately similar to 0501254A protein Tro alpha1 H,myeloma [H.sapiens]	Hs.293441
MMP13	Hs.2936
major histocompatibility complex, class II, DR beta 4	Hs.293934

Table 2: Candidate genes, Database mining

ESTs	Hs.279111
ESTs	Hs.279112
EST	Hs.279113
ESTs	Hs.279114
ESTs	Hs.279115
ESTs	Hs.279116
ESTs	Hs.279117
ESTs	Hs.279118
ESTs	Hs.279119
ESTs	Hs.279120
ESTs	Hs.279121
ESTs	Hs.279122
ESTs	Hs.279123
ESTs	Hs.279124
ESTs	Hs.279125
ESTs	Hs.279126
ESTs	Hs.279127
EST	Hs.279128
ESTs, Weakly similar to aconitase [H.sapiens]	Hs.279129
ESTs	Hs.279130
ESTs	Hs.279131
ESTs	Hs.279132
ESTs	Hs.279133
ESTs, Weakly similar to PYRG_HUMAN CTP SYNTHASE [H.sapiens]	Hs.279134
ESTs, Weakly similar to RIR1_HUMAN RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1 CHAIN [H.sapiens]	Hs.279135
ESTs	Hs.279136
ESTs	Hs.279137
ESTs	Hs.279138
ESTs	Hs.279139
ESTs	Hs.279140
ESTs	Hs.279141

Human MHC class III serum complement factor B, mRNA	Hs.294163
EST	Hs.294315
EST	Hs.294316
EST, Highly similar to Y196_HUMAN HYPOTHETICAL PROTEIN KIAA0196□ [H.sapiens]	Hs.295582
EST	Hs.295583
EST, Highly similar to ZN07_HUMAN ZINC FINGER PROTEIN 7 [H.sapiens]	Hs.295584
EST	Hs.295585
EST	Hs.295586
EST, Moderately similar to angiotensin converting enzyme [H.sapiens]	Hs.295595
EST	Hs.295621
EST	Hs.295622
EST, Moderately similar to RL13_HUMAN 60S RIBOSOMAL PROTEIN L13 [H.sapiens]	Hs.295629
EST	Hs.295724
EST	Hs.296064
EST, Moderately similar to IDS_HUMAN IDURONATE 2-SULFATASE PRECURSOR□ [H.sapiens]	Hs.296070
EST	Hs.296073
interleukin enhancer binding factor 1	Hs.296281
similar to rat integral membrane glycoprotein POM121	Hs.296429
Human histocompatibility antigen mrna clone phla-1	Hs.296476
immunoglobulin lambda-like polypeptide 3	Hs.296552
RFXANK	Hs.296776
	Hs.29826
	Hs.29871
MEKK1	Hs.298727
	Hs.30029
CD3e	Hs.3003
ESTs, Weakly similar to CA13_HUMAN COLLAGEN ALPHA 1(III) CHAIN PRECURSOR [H.sapiens]	Hs.300697
Homo sapiens clone BCSynL38 immunoglobulin lambda light chain variable region mRNA, partial cds	Hs.300865
FCGR3A	Hs.300983
Homo sapiens DP47 gene for immunoglobulin heavy chain, partial cds	Hs.301365
PMS2L9	Hs.301862

Table 2: Candidate genes, Database mining

EST	Hs.279142	CCR1	Hs.301921
ESTs	Hs.279143	FANCE	Hs.302003
ESTs	Hs.279144	interleukin 21	Hs.302014
ESTs	Hs.279145	interleukin 17E	Hs.302036
ESTs	Hs.279146		Hs.30446
EST	Hs.279147	EST	Hs.30709
ESTs	Hs.279148	EST	Hs.30731
ESTs	Hs.279149	MHC class II transactivator	Hs.3076
ESTs	Hs.279150	EST	Hs.30766
ESTs, Weakly similar to PUR2_HUMAN TRIFUNCTIONAL PURINE BIOSYNTHETIC PROTEIN ADENOSINE-3 [H.sapiens]	Hs.279151	EST	Hs.30793
ESTs	Hs.279152		Hs.30818
ESTs	Hs.279153	CD97	Hs.3107
ESTs	Hs.279154	RAR-beta2	Hs.31408
ESTs	Hs.279155	RECQL4	Hs.31442
ESTs	Hs.279156	XPC	Hs.320
ESTs	Hs.279157	ERK2	Hs.324473
ESTs	Hs.279158		Hs.32456
ESTs	Hs.279159	MSH6	Hs.3248
ESTs	Hs.279160	ribosomal protein L23-related	Hs.3254
ESTs, Weakly similar to IDHA_HUMAN ISOCITRATE DEHYDROGENASE [H.sapiens]	Hs.279161	PI3CG	Hs.32942
ESTs	Hs.279162	CSA (CKN1)	Hs.32967
ESTs	Hs.279163	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F	Hs.32981
ESTs	Hs.279164	BRCA2	Hs.34012
ESTs	Hs.279165	MEK1	Hs.3446
ESTs	Hs.279166	STRL33 (CXCR6)	Hs.34526
ESTs	Hs.279167	MBD4	Hs.35947
ESTs	Hs.279168	immunoglobulin (CD79A) binding protein 1	Hs.3631
EST	Hs.279169	CD7	Hs.36972
ESTs	Hs.279170	IFNA1	Hs.37026
ESTs	Hs.279171	PDGF-A	Hs.37040
EST	Hs.279172	immunoglobulin kappa variable 1-13	Hs.37089
ESTs	Hs.279174	DMC1	Hs.37181
ESTs	Hs.279175		Hs.37892
CD86	Hs.27954	Homo sapiens suppressor of variegation 3-9 (Drosophila) homolog (SUV39H) mRNA, and translated products.	Hs.37936
CGI-81 protein	Hs.279583	C8B	Hs.38069
ESTs	Hs.279821	MTH1 (NUDT1)	Hs.388
ESTs	Hs.279823	Adrenomedullin	Hs.394
ESTs, Weakly similar to IRE1_HUMAN IRON-RESPONSIVE ELEMENT BINDING PROTEIN 1 [H.sapiens]	Hs.279824		Hs.39441
ESTs	Hs.279825	CD66b	Hs.41
ESTs	Hs.279826	RAD50	Hs.41587

Table 2: Candidate genes, Database mining

MLH3	Hs.279843
TNFRSF14	Hs.279899
RPA4	Hs.283018
EST	Hs.283165
EST	Hs.283166
EST	Hs.283167
EST	Hs.283168
ESTs	Hs.283169
EST	Hs.283245
EST	Hs.283247
ESTs	Hs.283248
EST	Hs.283249
EST	Hs.283250
EST	Hs.283251
EST	Hs.283252
EST	Hs.283253
EST	Hs.283254
EST	Hs.283255
EST	Hs.283256
EST	Hs.283257
EST	Hs.283258
ESTs	Hs.283259
EST	Hs.283261
EST	Hs.283262
EST	Hs.283263
EST	Hs.283264
EST	Hs.283266
ESTs	Hs.283268
EST	Hs.283269
EST, Weakly similar to AF189011_1 ribonuclease III [H.sapiens]	Hs.283270
EST	Hs.283271
EST	Hs.283272
EST	Hs.283274
EST	Hs.283275
EST	Hs.283276
ESTs, Weakly similar to S32605 collagen alpha 3(VI) chain - mouse [M.musculus]	Hs.283392
ESTs	Hs.283433
ESTs	Hs.283434
ESTs	Hs.283438
ESTs	Hs.283442
ESTs	Hs.283443
ESTs	Hs.283456
ESTs	Hs.283457
ESTs, Weakly similar to similar to collagen [C.elegans]	Hs.283458

CD94	Hs.41682
HLJ1	Hs.41693
ESM1	Hs.41716
MSH3	Hs.42674
cAMP responsive element binding protein-like 1	Hs.42853
IKBKG	Hs.43505
Homo sapiens suppressor of white apricot homolog 2 (SWAP2), mRNA.	Hs.43543
LEU2	Hs.43628
Homo sapiens immunoglobulin lambda gene locus DNA, clone:288A10	Hs.43834
SIRT2	Hs.44017
	Hs.44087
TREM2	Hs.44234
serine/threonine kinase 19	Hs.444
	Hs.44512
	Hs.44628
	Hs.45063
LTC4 synthase	Hs.456
FUT2	Hs.46328
CCR6	Hs.46468
POLM	Hs.46964
EXO1 (HEX1)	Hs.47504
FEN1 (Dnase IV)	Hs.4756
	Hs.4863
golgin-165	Hs.4953
	Hs.50102
ATP-binding cassette, sub-family B (MDR/TAP), member 3	Hs.502
	Hs.5057
corneodesmosin	Hs.507
Histone H2 (H2AFP)	Hs.51011
CCNH	Hs.514
EST	Hs.5146
SMUG1	Hs.5212
ABH (ALKB)	Hs.54418
CCR5	Hs.54443
CD81	Hs.54457
TNFSF13	Hs.54673
PRPS1	Hs.56
	Hs.56156
	Hs.56265
killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 2	Hs.56328
EST	Hs.5656
	Hs.56845
MLH1	Hs.57301
testis specific basic protein	Hs.57692

Table 2: Candidate genes, Database mining

ESTs	Hs.283459
ESTs	Hs.283460
ESTs	Hs.283462
ESTs	Hs.283463
ESTs	Hs.283496
ESTs	Hs.283497
ESTs	Hs.283499
ESTs	Hs.283500
ESTs, Weakly similar to ORF YDL014w [S.cerevisiae]	Hs.283504
ESTs, Weakly similar to S09646 collagen alpha 2(VI) chain precursor, medium splice form - human [H.sapiens]	Hs.283505
ESTs	Hs.283608
CD42c	Hs.283743
tenascin XA	Hs.283750
immunoglobulin kappa variable 1D-8	Hs.283770
protocadherin gamma subfamily A, 2 (PCDHGA2)	Hs.283801
Homo sapiens mRNA; cDNA DKFZp762F0616 (from clone DKFZp762F0616)	Hs.283849
Homo sapiens clone bsmneg3-t7 immunoglobulin lambda light chain VJ region, (IGL) mRNA, partial cds	Hs.283876
Homo sapiens transgenic-JHD mouse #2357 immunoglobulin heavy chain variable region (IgG VH251) mRNA, partial cds	Hs.283878
Homo sapiens clone N97 immunoglobulin heavy chain variable region mRNA, partial cds	Hs.283882
Homo sapiens clone case06H1 immunoglobulin heavy chain variable region gene, partial cds	Hs.283924
Homo sapiens HSPC077 mRNA, partial cds	Hs.283929
Homo sapiens HSPC088 mRNA, partial cds	Hs.283931
Homo sapiens HSPC097 mRNA, partial cds	Hs.283933
Homo sapiens HSPC102 mRNA, partial cds	Hs.283934
Homo sapiens HSPC107 mRNA, partial cds	Hs.283935
CMKRL1	Hs.28408
FANCA	Hs.284153
Homo sapiens immunoglobulin mu chain antibody MO30 (IgM) mRNA, complete cds	Hs.284277
gamma-glutamyltransferase 1	Hs.284380

ESTs	Hs.57841
Human 6Ckine	Hs.57907
EST	Hs.5816
Homo sapiens cell growth regulatory with ring finger domain (CGR19) mRNA.	Hs.59106
ERCC1	Hs.59544
	Hs.61558
Homo sapiens GPI transamidase mRNA, complete cds.	Hs.62187
	Hs.62699
	Hs.63913
Homo sapiens chloride intracellular channel 3 (CLIC3), mRNA.	Hs.64746
FANCF	Hs.65328
	Hs.6544
interleukin 1 receptor-like 1	Hs.66
CD38	Hs.66052
	Hs.6607
RAD54L	Hs.66718
SCYA17 (CCL17)	Hs.66742
IL-12	Hs.673
Human IL-12 p40	Hs.674
LILRB4	Hs.67846
interleukin 5 receptor, alpha	Hs.68876
	Hs.6891
	Hs.69233
FUT1	Hs.69747
B-factor, properdin	Hs.69771
	Hs.70333
	Hs.71618
RAD1	Hs.7179
interleukin 19	Hs.71979

Table 2: Candidate genes, Database mining

putative human HLA class II associated protein I	Hs.285013
interleukin 13 receptor, alpha 1	Hs.285115
CDw131	Hs.285401
Homo sapiens VH2-D3.10-JH5b gene for immunoglobulin heavy chain variable region	Hs.287403
Homo sapiens cDNA: FLJ22546 fis, clone HSI00290	Hs.287697
Homo sapiens cDNA: FLJ23140 fis, clone LNG09065	Hs.287728
H.sapiens mRNA for HLA-C alpha chain (Cw*1701)	Hs.287811
Homo sapiens clone ASMneg1-b1 immunoglobulin lambda chain VJ region, (IGL) mRNA, partial cds	Hs.287815
Homo sapiens clone CPRF1-T2 immunoglobulin lambda chain VJ region, (IGL) mRNA, partial cds	Hs.287816
EST	Hs.287817
myelin protein zero-like 1	Hs.287832
immunoglobulin lambda-like polypeptide 1	Hs.288168
cathepsinB	Hs.288181
G18.2 protein	Hs.288316
ESTs	Hs.288403
EST	Hs.288431
Homo sapiens partial IGVH2 gene for immunoglobulin heavy chain V region, case 2, cell B 45	Hs.288553
polymeric immunoglobulin receptor	Hs.288579
Human immunoglobulin heavy chain variable region (V4-4) gene, partial cds	Hs.288711
Human immunoglobulin heavy chain variable region (V4-4b) gene, partial cds	Hs.289036
	Hs.28921
EST	Hs.289577
EST	Hs.289836
EST	Hs.289878
GSN	Hs.290070
EST, Weakly similar to unnamed protein product [H.sapiens]	Hs.290133
EST	Hs.290227
ESTs	Hs.290315
EST	Hs.290339
EST	Hs.290340
	Hs.29055
EST	Hs.291125
EST	Hs.291126
CD91= LRP	Hs.89137

MEK2	Hs.72241
IL-7	Hs.72927
STAT2	Hs.72988
CD42d	Hs.73734
MIF	Hs.73798
ECP	Hs.73839
CPN2	Hs.73858
MMP8	Hs.73862
HLA-G histocompatibility antigen, class I, G	Hs.73885
TNFRSF9	Hs.73895
IL-4	Hs.73917
HLA-DQB1	Hs.73931
RAG1	Hs.73958
LAG-3	Hs.74011
	Hs.7402
CD163	Hs.74076
immunoglobulin superfamily, member 2	Hs.74115
CD158b	Hs.74134
	Hs.7434
TCRa	Hs.74647
human immunodeficiency virus type I enhancer-binding protein 2	Hs.75063
MLN50	Hs.75080
lysyl hydroxylase (PLOD)	Hs.75093
TAK1	Hs.7510
Homo sapiens transcription factor 6-like 1 (mitochondrial transcription factor 1-like) (TCF6L1) mRNA.	Hs.75133
UBE2N (UBC13, BTG1)	Hs.75355
	Hs.75450
HSPA2	Hs.75452
CD151	Hs.75564
RELA	Hs.75569
CD122	Hs.75596
CD14	Hs.75627
nuclear factor erythroid 2 isoform f=basic leucine zipper protein {alternatively spliced	Hs.75643
C1QB	Hs.8986

Table 2: Candidate genes, Database mining

XPF (ERCC4)	Hs.89296
Carbonic anhydrase IV	Hs.89485
CETP	Hs.89538
RAD52	Hs.89571
GTF2H1	Hs.89578
Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide	Hs.897
transcript ch138	Hs.94881
	Hs.9578
IL-9	Hs.960
NFATC1	Hs.96149
OGG1	Hs.96398
	Hs.96499
NFKBIB	Hs.9731
XAB2 (HCNP)	Hs.9822
CD40	Hs652

superkiller viralicidic activity 2 (S. cerevisiae homolog)-like	Hs.89864
EST	Hs.90165
EST	Hs.90171
GTF2H3	Hs.90304
protein tyrosine kinase related sequence	Hs.90314
	Hs.90463
SGRF protein, Interleukin 23 p19 subunit	Hs.98309
XRCC1	Hs.98493
Homo sapiens mRNA for KIAA0543 protein, partial cds.	Hs.98507
	Hs.9893
DIR1 protein	Hs.99134
XRCC3	Hs.99742
Elastase(leukocyte)	Hs.99863
JAK3	Hs.99877

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

Example Clone	Offset on Acc Start	End	Accession Number	UniGene	Signif	Number Clones	Genbank Description
56D1	1521	1685	D00022	Hs.25	1.00E-84	1	for F1 beta subunit, complete
586E3	1227	1448	NM_001686	Hs.25	1.00E-89	1	ATP synthase, H+ transporting, mitochondrial
459F4	1484	2522	NM_002832	Hs.35	0	3	protein tyrosine phosphatase, non-receptor t
41A11	885	1128	D12614	Hs.36	1.00E-125	1	lymphotoxin (TNF-beta), complete
41G12	442	1149	D10202	Hs.46	0	1	for platelet-activating factor receptor,
98E12	1928	2652	NM_002835	Hs.62	0	1	protein tyrosine phosphatase, non-receptor t
170E1	473	1071	U13044	Hs.78	0	1	nuclear respiratory factor-2 subunit alpha mRNA, com
40C6	939	1357	D11086	Hs.84	0	1	*interleukin 2 receptor gamma chain*
521F9	283	1176	NM_000206	Hs.84	0	8	interleukin 2 receptor, gamma (severe combined
60A11	989	1399	L08069	Hs.94	0	2	heat shock protein, E. coli DnaJ homologue complete cd
520B9	545	1438	NM_001539	Hs.94	0	3	heat shock protein, DNAJ-like 2 (HSJ2), mRNA /
460H9	626	1104	NM_021127	Hs.96	0	1	phorbol-12-myristate-13-acetate-induced p
127G12	651	1223	NM_004906	Hs.119	0	2	Wilms' tumour 1-associating protein (KIAA0105
586A7	438	808	NM_000971	Hs.153	0	3	ribosomal protein L7 (RPL7), mRNA /cds=(10,756
99H12	2447	4044	NM_002600	Hs.188	0	2	phosphodiesterase 4B, cAMP-specific (dunce (
464D4	2317	2910	NM_002344	Hs.210	0	1	leukocyte tyrosine kinase (LTK), mRNA /cds=(17
464B3	10	385	NM_002515	Hs.214	1.00E-164	1	neuro-oncological ventral antigen 1 (NOVA1),
40A12	296	1153	L11695	Hs.220	0	1	activin receptor-like kinase (ALK-5) mRNA, complete
129A2	4138	4413	NM_000379	Hs.250	1.00E-155	1	xanthene dehydrogenase (XDH), mRNA
36B10	80	1475	AF068836	Hs.270	0	3	cytohesin binding protein HE mRNA, complete cd
45C11	58	1759	NM_004288	Hs.270	0	2	pleckstrin homology, Sec7 and coiled/coil dom
128C12	2555	3215	NM_000153	Hs.273	0	4	galactosylceramidase (Krabbe disease) (GALC)
67H2	259	1418	D23660	Hs.286	0	8	ribosomal protein, complete cds
151E6	624	1170	AF052124	Hs.313	0	1	clone 23810 osteopontin mRNA, complete cds /c
45A7	4	262	NM_000582	Hs.313	1.00E-136	1	secreted phosphoprotein 1 (osteopontin, bone
44C10	2288	2737	J03250	Hs.317	0	1	topoisomerase I mRNA, complete cds /cds=(211,2508) /
99H9	2867	3246	NM_001558	Hs.327	0	2	interleukin 10 receptor, alpha (IL10RA), mRNA
41B4	2867	3315	U00672	Hs.327	0	6	interleukin-10 receptor mRNA, complete
144E1	283	989	M26683	Hs.340	0	36	*interferon gamma treatment inducible* /cds=(14,1
41A12	1854	2590	X53961	Hs.347	0	1	lactoferrin /cds=(294,2429) /gb=X53961 /gi=
40F1	1377	1734	U95626	Hs.395	0	1	ccr2b (ccr2), ccr2a (ccr2), ccr5 (ccr5) and cc
463H4	55	434	NM_001459	Hs.428	0	1	fms-related tyrosine kinase 3 ligand (FLT3LG)
127E1	552	1048	NM_005180	Hs.431	0	1	murine leukemia viral (bmi-1) oncogene homolo
73G12	189	1963	NM_004024	Hs.460	0	17	activating transcription factor 3 (ATF3), ATF
524A4	1361	2136	NM_004168	Hs.469	0	2	succinate dehydrogenase complex, subunit A,
41C7	1554	2097	D10925	Hs.516	0	1	HM145 /cds=(22,1089) /gb=D10925 /gi=219862
588A2	48	163	NM_001032	Hs.539	1.00E-59	1	ribosomal protein S29 (RPS29), mRNA /cds=(30,2
177B4	1	1674	AF076465	Hs.550	2.00E-37	2	PhLOP2 mRNA, complete cds /cds=(5,358) /gb=AF
68G5	2	1454	M26383	Hs.624	0	17	monocyte-derived neutrophil-activating protein (M
45F10	1	1454	NM_000584	Hs.624	0	11	interleukin 8 (IL8), mRNA /cds=(74,373) /gb=N
59F11	59	1822	X68550	Hs.652	0	14	TRAP mRNA for ligand of CD40 /cds=(56,841) /gb=X6
471C9	3115	3776	NM_000492	Hs.663	0	1	cystic fibrosis transmembrane conductance re
68D1	228	866	M20137	Hs.694	0	3	interleukin 3 (IL-3) mRNA, complete cds, clone pcD-SR
49H3	42	665	NM_000588	Hs.694	0	1	interleukin 3 (colony-stimulating factor, mu
147H3	110	340	BF690338	Hs.695	1.00E-102	1	602186730T1 cDNA, 3' end /clone=IMAGE:4299006
483E4	310	846	NM_000942	Hs.699	0	1	peptidylprolyl isomerase B (cyclophilin B) (
522B12	349	755	NM_000788	Hs.709	0	2	deoxycytidine kinase (DCK), mRNA /cds=(159,94
331E5	1293	1470	J03634	Hs.727	9.00E-75	1	erythroid differentiation protein mRNA (EDF), comple
514D12	1164	1579	NM_004907	Hs.737	1.00E-169	3	immediate early protein (ETR101), mRNA /cds=(
73H7	1953	3017	AJ243425	Hs.738	0	8	EGR1 gene for early growth response protein 1 /

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

592A8	10	454	NM_003973	Hs.738	0	5	ribosomal protein L14 (RPL14), mRNA
519A1	116	1527	NM_000801	Hs.752	1.00E-163	2	FK506-binding protein 1A (12kD) (FKBP1A), mRN
109H11	1	1206	M60626	Hs.753	0	10	N-formylpeptide receptor (fMLP-R98) mRNA, complete
99C5	1	1175	NM_002029	Hs.753	0	25	formyl peptide receptor 1 (FPR1), mRNA
103C1	2285	2890	NM_002890	Hs.758	0	1	RAS p21 protein activator (GTPase activating p
41H4	3142	3332	NM_000419	Hs.785	1.00E-84	1	integrin, alpha 2b (platelet glycoprotein IIb
171D2	198	748	X54489	Hs.789	1.00E-132	2	melanoma growth stimulatory activity (MGSA)
458H7	2165	2818	NM_001656	Hs.792	0	1	ADP-ribosylation factor domain protein 1, 64
62B3	833	1241	M60278	Hs.799	0	2	heparin-binding EGF-like growth factor mRNA, complet
53G4	1299	2166	AK001364	Hs.808	0	6	FLJ10502 fis, clone NT2RP2000414, highly
597F3	1136	1797	NM_004966	Hs.808	0	2	heterogeneous nuclear ribonucleoprotein F (
143F7	575	985	M74525	Hs.811	0	3	HHR6B (yeast RAD 6 homologue) mRNA, complete
518H8	580	974	NM_003337	Hs.811	0	1	ubiquitin-conjugating enzyme E2B (RAD6 homol
45G8	277	833	NM_002121	Hs.814	0	1	major histocompatibility complex, class II,
41H11	719	1534	NM_005191	Hs.838	0	1	CD80 antigen (CD28 antigen ligand 1, B7-1 antig
41G1	117	557	U31120	Hs.845	0	1	interleukin-13 (IL-13) precursor gene, complete cds
75E1	693	862	J05272	Hs.850	2.00E-58	4	IMP dehydrogenase type 1 mRNA complete
129B11	3361	3883	L25851	Hs.851	0	1	integrin alpha E precursor, mRNA, complete cds
481E9	3361	3742	NM_002208	Hs.851	1.00E-173	1	integrin, alpha E (antigen CD103, human mucosa
71G7	1	1193	NM_000619	Hs.856	0	111	interferon, gamma (IFNG), mRNA /cds=(108,608)
75H5	1	1193	X13274	Hs.856	0	314	interferon IFN-gamma /cds=(108,608) /gb=X13
525B12	672	894	NM_002341	Hs.890	1.00E-121	1	lymphotoxin beta (TNF superfamily, member 3)
40E8	75	999	AL121985	Hs.901	0	6	DNA sequence RP11-404F10 on chromosome 1q2
48H4	680	933	NM_001778	Hs.901	1.00E-130	2	CD48 antigen (B-cell membrane protein) (CD48)
179G8	1652	2181	AL163285	Hs.926	0	1	chromosome 21 segment HS21C085
48G11	1049	2092	NM_002463	Hs.926	0	3	myxovirus (influenza) resistance 2, homolog o
110B12	209	1734	M32011	Hs.949	0	8	neutrophil oxidase factor (p67-phox) mRNA, complete
99C9	207	1733	NM_000433	Hs.949	0	11	neutrophil cytosolic factor 2 (65kD, chronic g
125D2	958	1645	NM_004645	Hs.966	0	1	coilin (COIL), mRNA /cds=(22,1752) /gb=NM_004
458C1	1649	2285	NM_006025	Hs.997	0	1	protease, serine, 22 (P11), mRNA /cds=(154,126
40H11	621	864	L26953	Hs.1010	1.00E-135	1	chromosomal protein mRNA, complete cds /cds=(7
116D10	513	858	NM_002932	Hs.1010	0	1	regulator of mitotic spindle assembly 1 (RMSA
40G11	1565	2151	M31452	Hs.1012	0	1	proline-rich protein (PRP) mRNA, complete
192A6	321	908	NM_000284	Hs.1023	0	1	pyruvate dehydrogenase (lipoamide) alpha 1 (
460H11	2158	2402	NM_004762	Hs.1050	2.00E-91	1	pleckstrin homology, Sec7 and coiled/coil dom
41F12	291	565	M57888	Hs.1051	1.00E-112	1	(clone lambda B34) cytotoxic T-lymphocyte-associate
41A5	1311	1852	M55654	Hs.1100	0	1	TATA-binding protein mRNA, complete
461D7	999	1277	NM_002698	Hs.1101	1.00E-92	1	POU domain, class 2, transcription factor 2 (P
597H9	1083	1224	NM_000660	Hs.1103	3.00E-75	1	transforming growth factor, beta 1 (TGFB1), mR
40B5	1433	2010	X02812	Hs.1103	0	1	transforming growth factor-beta (TGF-beta)
106A10	1977	2294	M73047	Hs.1117	1.00E-176	1	tripeptidyl peptidase II mRNA, complete cds /c
165E8	4273	4582	NM_003291	Hs.1117	1.00E-173	1	tripeptidyl peptidase II (TPP2), mRNA /cds=(23
63G12	1114	2339	D49728	Hs.1119	0	7	NAK1 mRNA for DNA binding protein, complete
45B10	1317	1857	NM_002135	Hs.1119	0	1	nuclear receptor subfamily 4, group A, member
37H3	568	783	M24069	Hs.1139	1.00E-119	1	DNA-binding protein A (dbpA) gene, 3' end
476F9	209	608	NM_000174	Hs.1144	0	1	glycoprotein IX (platelet) (GP9), mRNA /cds=(
43A10	1105	1357	U15085	Hs.1162	3.00E-41	1	HLA-DMB mRNA, complete cds
139D6	1345	1680	L11329	Hs.1183	1.00E-102	1	protein tyrosine phosphatase (PAC-1) mRNA, co
134B12	1233	1675	NM_004418	Hs.1183	0	1	dual specificity phosphatase 2 (DUSP2), mRNA
58F1	17	341	NM_002157	Hs.1197	0	1	heat shock 10kD protein 1 (chaperonin 10) (HSP
158G5	20	341	U07550	Hs.1197	1.00E-180	2	chaperonin 10 mRNA, complete cds
167C8	813	1453	NM_000022	Hs.1217	0	4	adenosine deaminase (ADA), mRNA /cds=(95,1186
179H1	730	1452	X02994	Hs.1217	0	6	adenosine deaminase (adenosine aminohydrola
40E10	594	792	M38690	Hs.1244	1.00E-109	1	CD9 antigen mRNA, complete cds

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

41C5	1280	1438	AK024951	Hs.1279	2.00E-80	1	FLJ21298 fis, clone COL02040, highly sim
40E3	1002	1735	NM_000065	Hs.1282	0	1	complement component 6 (C6) mRNA /cd
40A11	1638	1821	K02766	Hs.1290	3.00E-98	1	complement component C9 mRNA, complete
40B12	4639	5215	NM_007289	Hs.1298	0	1	membrane metallo-endopeptidase (neutral end
41G2	1576	1870	M28825	Hs.1309	1.00E-115	1	thymocyte antigen CD1a mRNA, complete cds
41F8	1171	1551	AX023365	Hs.1349	0	1	Sequence 36 from Patent WO0006605
40E1	673	1147	M30142	Hs.1369	0	1	decay-accelerating factor mRNA, complete cds
118B12	1129	1719	NM_000574	Hs.1369	0	1	decay accelerating factor for complement (CD5
75F8	830	2979	NM_000399	Hs.1395	0	48	early growth response 2 (Krox-20 (Drosophila)
41F11	973	1428	M15059	Hs.1416	0	1	Fc-epsilon receptor (IgE receptor) mRNA, complete cd
110G12	1931	2071	AL031729	Hs.1422	2.00E-70	1	DNA seq RP1-159A19 on chromosome 1p36
113D10	1718	2066	NM_005248	Hs.1422	6.00E-76	2	Gardner-Rasheed feline sarcoma viral (v-fgr)
477C2	3292	3842	NM_000152	Hs.1437	0	1	glucosidase, alpha; acid (Pompe disease, glyc
124D1	795	1127	NM_000167	Hs.1466	0	1	glycerol kinase (GK), mRNA /cds=(66,1640) /gb
41B9	2231	2447	J03171	Hs.1513	1.00E-108	1	interferon-alpha receptor (HuIFN-alpha-Rec) mRNA,
99F7	927	1889	NM_014882	Hs.1528	0	2	KIAA0053 gene product (KIAA0053), mRNA /cds=(
469G9	1220	1507	NM_005082	Hs.1579	1.00E-117	1	zinc finger protein 147 (estrogen-responsive
195B7	190	1801	BC002971	Hs.1600	0	3	clone IMAGE:3543711, mRNA, partial cds /cds=
195F10	3676	3856	NM_000110	Hs.1602	1.00E-85	1	dihydropyrimidine dehydrogenase (DPYD), mRN
129E7	648	1827	L08176	Hs.1652	0	2	Epstein-Barr virus induced G-protein coupled recepto
478H5	1839	2050	NM_002056	Hs.1674	7.00E-79	1	glutamine-fructose-6-phosphate transaminas
39H1	436	865	L35249	Hs.1697	0	1	vacuolar H+-ATPase Mr 56,000 subunit (HO57) mR
183H8	972	1183	NM_001693	Hs.1697	1.00E-106	1	ATPase, H+ transporting, lysosomal (vacuolar
481A4	1594	1785	NM_001420	Hs.1701	2.00E-79	1	ELAV (embryonic lethal, abnormal vision, Dros
40B3	3846	4009	L39064	Hs.1702	4.00E-70	1	interleukin 9 receptor precursor (IL9R) gene,
176G8	1033	1400	NM_006084	Hs.1706	0	1	interferon-stimulated transcription factor
589C11	1	1347	NM_005998	Hs.1708	0	2	chaperonin containing TCP1, subunit 3 (gamma)
70H5	1	494	X74801	Hs.1708	0	1	Cctg mRNA for chaperonin /cds=(0,1634) /gb=X7480
460C12	3310	3809	NM_012089	Hs.1710	0	1	ATP-binding cassette, sub-family B (MDR/TAP),
41D5	484	1862	M28983	Hs.1722	0	3	interleukin 1 alpha (IL 1) mRNA, complete cds /
119E8	493	904	NM_000575	Hs.1722	1.00E-151	2	interleukin 1, alpha (IL1A), mRNA /cds=(36,851
479E11	5	268	NM_000417	Hs.1724	1.00E-145	1	interleukin 2 receptor, alpha (IL2RA), mRNA /
62C8	85	1887	X01057	Hs.1724	0	2	interleukin-2 receptor /cds=(180,998) /gb=X
466A3	2166	2675	NM_000889	Hs.1741	0	1	integrin, beta 7 (ITGB7), mRNA /cds=(151,2547)
107A4	4960	5610	L33075	Hs.1742	0	1	ras GTPase-activating-like protein (IQGAP1)
189A5	4318	7450	NM_003870	Hs.1742	0	3	IQ motif containing GTPase activating protein
597D1	1230	1737	NM_005356	Hs.1765	1.00E-127	5	lymphocyte-specific protein tyrosine kinase
41C10	1057	1602	J04142	Hs.1799	0	1	(lambda-gt11ht-5) MHC class I antigen-like gl
104H1	1854	2023	L06175	Hs.1845	4.00E-54	1	P5-1 mRNA, complete cds /cds=(304,735) /gb=L06
98F7	34	2041	NM_006674	Hs.1845	4.00E-63	5	MHC class I region ORF (P5-1), /cds=(304,735) /
104F1	1390	1756	NM_002436	Hs.1861	0	2	membrane protein, palmitoylated 1 (55kD) (MPP
171F7	1760	2192	M55284	Hs.1880	0	1	protein kinase C-L (PRKCL) mRNA, complete cds
134B2	123	1182	NM_002727	Hs.1908	0	10	proteoglycan 1, secretory granule (PRG1), mRN
61C11	126	902	X17042	Hs.1908	0	11	hematopoetic proteoglycan core protein /cds
458G1	1	475	NM_001885	Hs.1940	0	1	crystallin, alpha B (CRYAB), mRNA
520E10	71	343	NM_001024	Hs.1948	1.00E-142	3	ribosomal protein S21 (RPS21), mRNA
459D6	2435	3055	NM_001761	Hs.1973	0	1	cyclin F (CCNF), mRNA /cds=(43,2403)
41H3	184	1620	NM_006139	Hs.1987	0	2	CD28 antigen (Tp44) (CD28), mRNA /cds=(222,884
71C5	721	1329	NM_000639	Hs.2007	0	2	tumor necrosis factor (ligand) superfamily, m
73C1	721	1603	X89102	Hs.2007	0	8	fasligand /cds=(157,1002)
135G3	940	1352	NM_002852	Hs.2050	6.00E-96	1	pentaxin-related gene, rapidly induced by IL
44A10	1562	1748	M58028	Hs.2055	7.00E-69	1	ubiquitin-activating enzyme E1 (UBE1) mRNA, complete

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

155G5	973	2207	AL133415	Hs.2064	0	7	DNA sequence from clone RP11-124N14 on chromosome 10.
599H7	48	3022	AK025306	Hs.2083	0	12	cDNA: FLJ21653 fis, clone COL08586,
71H1	1598	2163	NM_004419	Hs.2128	0	5	dual specificity phosphatase 5 (DUSP5), mRNA
69H7	1595	2161	U15932	Hs.2128	0	11	dual-specificity protein phosphatase mRNA, complete
458C4	1928	2356	NM_005658	Hs.2134	0	1	TNF receptor-associated factor 1 (TRAF1), mRN
192E11	6	414	NM_002704	Hs.2164	0	1	pro-platelet basic protein (includes platele
40D12	1935	2645	M58597	Hs.2173	0	2	ELAM-1 ligand fucosyltransferase (ELFT) mRNA, comple
40E5	2834	3024	M59820	Hs.2175	1.00E-104	1	granulocyte colony-stimulating factor receptor (CSF
482D8	2521	2943	NM_000760	Hs.2175	0	2	colony stimulating factor 3 receptor (granuloc
60H6	918	1723	AF119850	Hs.2186	0	6	PRO1608 mRNA, complete cds /cds=(1221,2174) /
597F11	99	1267	NM_001404	Hs.2186	0	29	eukaryotic translation elongation factor 1 g
595G4	6	570	L40410	Hs.2210	0	1	thyroid receptor interactor (TRIP3) mRNA, 3'
41H12	970	1353	X03656	Hs.2233	0	1	granulocyte colony-stimulating factor (G-C
461A9	287	730	Z29067	Hs.2236	0	1	H.sapiens nek3 mRNA for protein kinase
493E11	212	608	NM_000879	Hs.2247	1.00E-141	2	interleukin 5 (colony-stimulating factor, eo
150B5	363	815	X04688	Hs.2247	0	1	T-cell replacing factor (interleukin-5) /cd
461E12	255	342	NM_001565	Hs.2248	8.00E-34	1	small inducible cytokine subfamily B (Cys-X-C
129A8	1790	1970	NM_002309	Hs.2250	2.00E-94	1	leukemia inhibitory factor (cholinergic diff
40G10	2152	2560	X04481	Hs.2253	0	1	complement component C2 /cds=(36,2294) /gb=X
479A2	95	610	NM_000073	Hs.2259	0	2	CD3G antigen, gamma polypeptide (TiT3 complex
592G6	783	1163	NM_002950	Hs.2280	0	2	ribophorin I (RPN1), mRNA /cds=(137,1960) /gb
459G11	673	1316	NM_004931	Hs.2299	0	1	CD8 antigen, beta polypeptide 1 (p37) (CD8B1),
129B8	1159	1316	X13444	Hs.2299	1.00E-74	1	CD8 beta-chain glycoprotein (CD8 beta.1) /cd
467F12	2928	3239	NM_000346	Hs.2316	3.00E-85	1	SRY (sex determining region Y)-box 9 (campomeli
44A6	1506	1629	U23028	Hs.2437	7.00E-62	1	eukaryotic initiation factor 2B-epsilon mRNA, partia
127B8	1814	2405	NM_003816	Hs.2442	0	1	a disintegrin and metalloproteinase domain 9
36G6	1361	2019	D13645	Hs.2471	0	2	KIAA0020 gene, complete cds /cds=(418,1944)
458D6	396	961	NM_021966	Hs.2484	0	1	T-cell leukemia/lymphoma 1A (TCL1A), mRNA /c
124G1	966	1473	NM_005565	Hs.2488	0	1	lymphocyte cytosolic protein 2 (SH2 domain-con
107A6	1962	2031	U20158	Hs.2488	2.00E-22	1	76 kDa tyrosine phosphoprotein SLP-76 mRNA, complete
592E12	2175	2458	NM_002741	Hs.2499	1.00E-158	1	protein kinase C-like 1 (PRKCL1), mRNA /cds=(8
106A11	1455	2219	U34252	Hs.2533	0	2	gamma-aminobutyraldehyde dehydrogenase mRNA, compl
40F8	2201	2694	NM_003032	Hs.2554	0	1	sialyltransferase 1 (beta-galactoside alpha-
460G6	565	2052	NM_002094	Hs.2707	0	2	G1 to S phase transition 1 mRNA
60G5	35	184	X92518	Hs.2726	7.00E-27	2	HMGI-C protein /cds=UNKNOWN
461F10	1034	1520	NM_002145	Hs.2733	0	2	homeo box B2 (HOXB2), mRNA
69G2	408	1369	AK026515	Hs.2795	0	4	FLJ22862 fis, clone KAT01966, highly sim
71D8	13	541	NM_005566	Hs.2795	0	1	lactate dehydrogenase A (LDHA), mRNA /cds=(97
40H12	4119	4807	NM_002310	Hs.2798	0	1	leukemia inhibitory factor receptor (LIFR) mR
189C12	696	1287	NM_006196	Hs.2853	0	2	poly(rC)-binding protein 1 (PCBP1), mRNA /cds
111E8	1298	1938	NM_003566	Hs.2864	0	1	early endosome antigen 1, 162kD (EEA1), mRNA /
127F12	34	248	NM_001033	Hs.2934	1.00E-109	1	ribonucleotide reductase M1 polypeptide (RRM
74G6	11	241	AK023088	Hs.2953	1.00E-128	38	FLJ13026 fis, clone NT2RP3000968, modera
128D8	178	518	NM_000117	Hs.2985	1.00E-173	1	emerin (Emery-Dreifuss muscular dystrophy) (
169G7	2406	3112	AL136593	Hs.3059	0	1	DKFZp761K102 (from clone DKFZp761K1
193A3	2405	3017	NM_016451	Hs.3059	0	5	coatomer protein complex, subunit beta (COPB)
53F12	486	1007	L11066	Hs.3069	0	3	sequence /cds=UNKNOWN /gb=L11066 /gi=307322 /u
71E8	1623	2131	NM_004134	Hs.3069	0	2	heat shock 70kD protein 9B (mortalin-2) (HSPA9
458A5	2236	2874	NM_014877	Hs.3085	0	1	KIAA0054 gene product; Helicase (KIAA0054), m
69E8	1752	1916	D31884	Hs.3094	7.00E-68	1	KIAA0063 gene, complete cds /cds=(279,887) /
66B3	251	1590	D32053	Hs.3100	0	2	for Lysyl tRNA Synthetase, complete cds /
458E1	1645	1964	NM_001666	Hs.3109	1.00E-178	1	Rho GTPase activating protein 4 (ARHGAP4), mRN

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

331D8	2882	3585	U26710	Hs.3144	0	1	cbl-b mRNA, complete cds /cds=(322,3270) /gb=U26710
73D9	1	613	AL031736	Hs.3195	0	18	DNA sequence clone 738P11 on chromosome 1q24.1-2
58B1	1	607	NM_002995	Hs.3195	0	17	small inducible cytokine subfamily C, member
98F11	145	588	NM_003172	Hs.3196	0	1	surfeit 1 (SURF1), mRNA /cds=(14,916) /gb=NM_
124E9	1258	2414	NM_007318	Hs.3260	0	2	presenilin 1 (Alzheimer disease 3) (PSEN1), tr
64G7	1040	1569	NM_002155	Hs.3268	0	1	heat shock 70kD protein 6 (HSP70B') (HSPA6), mR
36D4	1116	1917	X51757	Hs.3268	0	4	heat-shock protein HSP70B' gene /cds=(0,1931) /gb=X5
39H11	1	507	BE895166	Hs.3297	1.00E-152	4	601436095F1 cDNA, 5' end /clone=IMAGE:3921239
103G4	16	540	NM_002954	Hs.3297	0	4	ribosomal protein S27a (RPS27A), mRNA /cds=(3
127H7	1391	1806	AB037752	Hs.3355	0	1	mRNA for KIAA1331 protein, partial cds /cds=(0
107D3	1932	2517	AK027064	Hs.3382	0	1	FLJ23411 fis, clone HEP20452, highly sim
121B3	1270	3667	NM_005134	Hs.3382	0	4	protein phosphatase 4, regulatory subunit 1 (
58H1	104	573	NM_001122	Hs.3416	0	6	adipose differentiation-related protein (AD
75G1	104	1314	X97324	Hs.3416	0	16	adipophilin /cds=(0,1313) /gb=X97324 /
182A4	147	334	NM_001867	Hs.3462	1.00E-102	1	cytochrome c oxidase subunit VIIc (COX7C), mRN
134D7	36	270	NM_001025	Hs.3463	1.00E-127	3	ribosomal protein S23 (RPS23), mRNA /cds=(13,4
192B10	129	1135	AL357536	Hs.3576	0	3	mRNA full length insert cDNA clone EUROIMAGE 37
112G12	56	687	NM_003001	Hs.3577	0	1	succinate dehydrogenase complex, subunit C,
526H6	143	537	BF666961	Hs.3585	0	1	602121608F1 cDNA, 5' end /clone=IMAGE:4278768
599F10	2098	2351	NM_004834	Hs.3628	1.00E-118	2	mitogen-activated protein kinase kinase kina
594F1	239	1321	NM_001551	Hs.3631	0	4	immunoglobulin (CD79A) binding protein 1 (IG
463E7	911	1033	AL359940	Hs.3640	1.00E-63	1	mRNA; cDNA DKFZp762P1915 (from clone DKFZp762P
182A9	657	1179	AL050268	Hs.3642	0	2	mRNA; cDNA DKFZp564B163 (from clone DKFZp564B1
38B4	257	568	AB034205	Hs.3688	1.00E-151	3	for cisplatin resistance-associated ove
185H6	769	995	NM_006003	Hs.3712	2.00E-88	1	ubiquinol-cytochrome c reductase, Rieske iro
587A1	716	1609	NM_006007	Hs.3776	0	2	zinc finger protein 216 (ZNF216), mRNA /cds=(2
473B5	46	531	NM_021633	Hs.3826	0	1	kelch-like protein C3IP1 (C3IP1), mRNA /cds=(
194G5	2456	2984	AB002366	Hs.3852	0	1	mRNA for KIAA0368 gene, partial cds /cds=(0,4327) /gb
589B4	526	1337	NM_000310	Hs.3873	0	3	palmitoyl-protein thioesterase 1 (ceroid-lip
515A10	1618	2130	NM_002267	Hs.3886	0	1	karyopherin alpha 3 (importin alpha 4) (KPNA3)
186A8	1160	1632	NM_002807	Hs.3887	0	1	proteasome (prosome, macropain) 26S subunit,
102F7	4226	4531	AB023163	Hs.4014	1.00E-158	1	for KIAA0946 protein, partial cds /cds=(0
50B8	1	166	AL117595	Hs.4055	3.00E-89	2	cDNA DKFZp564C2063 (from clone DKFZp564
473A10	1064	1709	NM_006582	Hs.4069	0	1	glucocorticoid modulatory element binding pr
524A12	2863	3386	AL136105	Hs.4082	0	1	DNA sequence from clone RP4-670F13 on chromosome 1q42
525E1	521	974	BC002435	Hs.4096	0	1	clone IMAGE:3346451, mRNA, partial cds /cds=
163G12	1130	1630	X52882	Hs.4112	0	6	t-complex polypeptide 1 gene /cds=(21,1691) /gb=X528
176A7	515	892	BC000687	Hs.4147	0	1	translocating chain-associating membrane p
185B5	3480	3707	AB023216	Hs.4278	1.00E-86	1	mRNA for KIAA0999 protein, partial cds /cds=(0
154E12	1731	2531	AF079566	Hs.4311	0	2	ubiquitin-like protein activating enzyme (UB
331C9	1595	1966	AF067008	Hs.4747	0	1	dyskerin (DKC1) mRNA, complete cds /cds=(60,16
182C8	1676	1966	NM_001363	Hs.4747	1.00E-148	2	dyskeratosis congenita 1, dyskerin (DKC1), mR
178C4	1623	2162	AL136610	Hs.4750	0	3	mRNA; cDNA DKFZp564K0822 (from clone DKFZp564K
107F9	3857	4266	AB032976	Hs.4779	0	1	for KIAA1150 protein, partial cds /cds=(0
191C11	1945	2618	AF240468	Hs.4788	0	3	nicastrin mRNA, complete cds /cds=(142,2271)
143G11	869	2076	AK022974	Hs.4859	0	2	FLJ12912 fis, clone NT2RP2004476, highly
127H11	977	1666	NM_020307	Hs.4859	0	1	cyclin L ania-6a (LOC57018), mRNA /cds=(54,163
479A11	215	544	AK001942	Hs.4863	1.00E-173	1	cDNA FLJ11080 fis, clone PLACE1005181 /cds=UN

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

73C5	2314	2851	AF105366	Hs.4876	0	1	K-Cl cotransporter KCC3a mRNA, alternatively
525F9	1059	1764	NM_006513	Hs.4888	0	3	seryl-tRNA synthetase (SARS), mRNA /cds=(75,1
114D8	931	1061	Z24724	Hs.4934	4.00E-52	1	H.sapiens polyA site DNA /cds=UNKNOWN /gb=Z24724 /gi=50503
587C10	1104	1343	NM_006787	Hs.4943	3.00E-94	1	hepatocellular carcinoma associated protein;
174F12	1749	2291	NM_018107	Hs.4997	0	3	hypothetical protein FLJ10482 (FLJ10482), mR
514C11	899	1489	AK021776	Hs.5019	0	1	cDNA FLJ11714 fis, clone HEMBA1005219, weakly
126H9	25	397	BE379724	Hs.5027	1.00E-118	1	601159415T1 cDNA, 3' end /clone=IMAGE:3511107
599B5	801	970	NM_017840	Hs.5080	5.00E-73	1	hypothetical protein FLJ20484 (FLJ20484), mR
47E5	4	720	AL034553	Hs.5085	0	2	DNA sequence from clone 914P20 on chromosome 20q13.13
122C11	492	860	NM_003859	Hs.5085	0	1	dolichyl-phosphate mannosyltransferase pol
116H6	1644	2902	NM_014868	Hs.5094	1.00E-102	2	ring finger protein 10 (RNF10), mRNA /cds=(698,
187G7	700	1268	NM_004710	Hs.5097	0	1	synaptogyrin 2 (SYNGR2), mRNA /cds=(29,703) /
174G3	240	500	NM_003746	Hs.5120	1.00E-144	4	dynein, cytoplasmic, light polypeptide (PIN)
145B6	199	695	BE539096	Hs.5122	1.00E-165	2	601061641F1 cDNA, 5' end /clone=IMAGE:3447850
486C1	1	529	BG028906	Hs.5122	0	2	602293015F1 cDNA, 5' end /clone=IMAGE:4387778
69F6	62	455	BF307213	Hs.5174	0	1	601891365F1 cDNA, 5' end /clone=IMAGE:4136752
583F4	82	477	NM_001021	Hs.5174	0	1	ribosomal protein S17 (RPS17), mRNA /cds=(25,4
74C4	1955	2373	AK025367	Hs.5181	1.00E-179	1	FLJ21714 fis, clone COL10256, highly sim
73E12	702	987	AL109840	Hs.5184	1.00E-161	1	DNA sequence from clone RP4-543J19 on chromosome 20 C
180G4	26	639	NM_002212	Hs.5215	0	2	integrin beta 4 binding protein (ITGB4BP), mRN
98F1	17	636	NM_014165	Hs.5232	0	5	HSPC125 protein (HSPC125), mRNA /cds=(79,606)
525A8	479	992	NM_006698	Hs.5300	0	1	bladder cancer associated protein (BLCAP), mR
99C1	19	507	NM_003333	Hs.5308	0	3	ubiquitin A-52 residue ribosomal protein fusi
172D11	714	1805	NM_005721	Hs.5321	0	3	ARP3 (actin-related protein 3, yeast) homolog
591F6	475	970	NM_015702	Hs.5324	0	1	hypothetical protein (CL25022), mRNA /cds=(1
68H8	724	1190	NM_014106	Hs.5327	0	2	PRO1914 protein (PRO1914), mRNA /cds=(1222,14
194D12	2128	2499	AB018305	Hs.5378	0	1	mRNA for KIAA0762 protein, partial cds /cds=(0
501G11	823	1322	NM_020122	Hs.5392	0	3	potassium channel modulatory factor (DKFZP434
74B4	502	1257	AF008442	Hs.5409	0	7	RNA polymerase I subunit hRPA39 mRNA, complete
134H7	543	916	NM_004875	Hs.5409	0	1	RNA polymerase I subunit (RPA40), mRNA /cds=(2
168A3	1909	2379	AF090891	Hs.5437	0	1	clone HQ0105 PRO0105 mRNA, complete cds /cds=(
145C10	2375	2564	AF016270	Hs.5464	1.00E-104	2	thyroid hormone receptor coactivating protein
587H7	1857	2563	NM_006696	Hs.5464	0	4	thyroid hormone receptor coactivating protein
183D10	1199	1347	NM_006495	Hs.5509	9.00E-40	1	ecotropic viral integration site 2B (EVI2B), m
181D7	1385	1752	AK002173	Hs.5518	0	1	cDNA FLJ11311 fis, clone PLACE1010102 /cds=UNK
173B1	1	642	NM_003315	Hs.5542	0	2	tetratricopeptide repeat domain 2 (TTC2), mRN
120F8	1782	2430	AF157323	Hs.5548	0	2	p45SKP2-like protein mRNA, complete cds /cds=
464H2	46	357	NM_000998	Hs.5566	1.00E-163	2	ribosomal protein L37a (RPL37A), mRNA /cds=(1
75F5	1252	2194	AK027192	Hs.5615	0	9	FLJ23539 fis, clone LNG08101, highly sim
56E8	27	205	AI570531	Hs.5637	2.00E-95	1	tm77g04.x1 cDNA, 3' end /clone=IMAGE:2164182
524G2	2	926	NM_006098	Hs.5662	0	9	guanine nucleotide binding protein (G protein
39F6	2311	2902	AB014579	Hs.5734	0	1	for KIAA0679 protein, partial cds /cds=(0
587G2	2883	4606	NM_012215	Hs.5734	0	11	meningioma expressed antigen 5 (hyaluronidase
469E5	5041	5393	NM_014864	Hs.5737	3.00E-75	2	KIAA0475 gene product (KIAA0475), mRNA /cds=(
120H3	1022	1553	NM_016230	Hs.5741	0	1	flavohemoprotein b5+b5R (LOC51167), mRNA /cd
63H8	1049	1507	AK025729	Hs.5798	0	1	FLJ22076 fis, clone HEP12479, highly sim
590D9	1015	1470	NM_015946	Hs.5798	0	1	pelota (Drosophila) homolog (PELO), mRNA /cds
102E3	665	1027	AK000474	Hs.5811	0	1	FLJ20467 fis, clone KAT06638 /cds=(360,77

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

187E5	665	1028	NM_017835	Hs.5811	0	1	chromosome 21 open reading frame 59 (C21ORF59),
39F9	1402	1728	AK025773	Hs.5822	0	3	FLJ22120 fis, clone HEP18874 /cds=UNKNOW
39E12	1064	1843	AF208844	Hs.5862	0	1	BM-002 mRNA, complete cds /cds=(39,296) /gb=A
173H9	906	1684	NM_016090	Hs.5887	0	2	RNA.binding motif protein 7 (LOC51120), mRNA /
120E8	1702	2055	NM_012179	Hs.5912	1.00E-146	1	F-box only protein 7 (FBXO7), mRNA /cds=(205,17
195D1	1309	2656	AK025620	Hs.5985	0	8	cDNA: FLJ21967 fis, clone HEP05652, highly sim
116A6	1451	2073	AK024941	Hs.6019	0	1	cDNA: FLJ21288 fis, clone COL01927 /cds=UNKNOW
113F9	1232	1598	NM_002896	Hs.6106	1.00E-126	1	RNA binding motif protein 4 (RBM4), mRNA /cds=(
520H1	563	1007	NM_018285	Hs.6118	0	2	hypothetical protein FLJ10968 (FLJ10968), mR
180H12	5224	5568	AF315591	Hs.6151	1.00E-135	1	Pumilio 2 (PUMH2) mRNA, complete cds /cds=(23,3
185A7	612	1558	NM_016001	Hs.6153	0	6	CGI-48 protein (LOC51096), mRNA /cds=(107,167
595G2	3207	4752	Z97056	Hs.6179	0	10	DNA seq from clone RP3-434P1 on chromosome 22
592B11	234	4611	AI745230	Hs.6187	1.00E-130	6	wg10e05.x1 cDNA, 3' end /clone=IMAGE:2364704
590F2	994	1625	NM_004517	Hs.6196	0	3	integrin-linked kinase (ILK), mRNA /cds=(156,
188A3	1550	2929	M61906	Hs.6241	0	3	P13-kinase associated p85 mRNA sequence
103C12	502	1129	AF246238	Hs.6289	0	1	HT027 mRNA, complete cds /cds=(260,784) /gb=A
100C2	804	1111	AK024539	Hs.6289	1.00E-122	1	FLJ20886 fis, clone ADKA03257 /cds=(359,
480A11	1149	1242	AB032977	Hs.6298	1.00E-46	1	mRNA for KIAA1151 protein, partial cds /cds=(0
473C8	3944	4149	NM_014859	Hs.6336	1.00E-106	1	KIAA0672 gene product (KIAA0672), mRNA /cds=(
125A10	1293	1766	NM_006791	Hs.6353	0	1	MORF-related gene 15 (MRG15), mRNA /cds=(131,1
182F5	143	2118	NM_018471	Hs.6375	0	3	uncharacterized hypothalamus protein HT010
587E8	398	2287	NM_016289	Hs.6406	0	7	MO25 protein (LOC51719), mRNA /cds=(53,1078)
135C3	2519	3084	AF130110	Hs.6456	0	2	clone FLB6303 PRO1633 mRNA, complete cds /cds=
178B5	1744	2425	AL117352	Hs.6523	0	2	DNA seq from clone RP5-876B10 on chromosome 1q42
522F10	2392	2591	NM_001183	Hs.6551	1.00E-110	2	ATPase, H+ transporting, lysosomal (vacuolar
595C4	1676	2197	NM_021008	Hs.6574	0	4	suppressin (nuclear deformed epidermal autor
481F3	745	904	AL117565	Hs.6607	9.00E-82	1	mRNA; cDNA DKFZp566F164 (from clone DKFZp566F1
124A3	1046	1575	NM_017792	Hs.6631	0	1	hypothetical protein FLJ20373 (FLJ20373), mR
177F11	1966	2281	AB046844	Hs.6639	1.00E-152	1	for KIAA1624 protein, partial cds /cds=(0
521G7	4600	5210	NM_014856	Hs.6684	0	2	KIAA0476 gene product (KIAA0476), mRNA /cds=(
54C6	265	756	AB037801	Hs.6685	0	1	for KIAA1380 protein, partial cds /cds=(0
75F7	95	3507	AB014560	Hs.6727	0	4	for KIAA0660 protein, complete cds /cds=(
477H12	2	457	BF976590	Hs.6749	0	1	602244267F1 cDNA, 5' end /clone=IMAGE:4335353
60A1	1028	1307	AB026908	Hs.6790	1.00E-155	1	for microvascular endothelial differenti
100G9	341	454	BE875609	Hs.6820	2.00E-58	1	601487048F1 cDNA, 5' end /clone=IMAGE:3889762
184F7	1259	1633	AF056717	Hs.6856	0	5	ash2l2 (ASH2L2) mRNA, complete cds /cds=(295,1
195E7	1250	1711	NM_004674	Hs.6856	0	3	ash2 (absent, small, or homeotic, Drosophila,
135F11	328	600	NM_020188	Hs.6879	1.00E-151	1	DC13 protein (DC13), mRNA /cds=(175,414) /gb=
172G2	1477	1782	NM_015530	Hs.6880	1.00E-169	1	DKFZP434D156 protein (DKFZP434D156), mRNA /c
483G5	3712	3947	AL031681	Hs.6891	3.00E-72	1	DNA sequence from clone 862K6 on chromosome 20q12-13.1
184B1	1	622	AF006086	Hs.6895	0	3	Arp2/3 protein complex subunit p21-Arc (ARC21
599C12	1	622	NM_005719	Hs.6895	0	24	actin related protein 2/3 complex, subunit 3 (
43A1	2111	2312	AF037204	Hs.6900	9.00E-78	1	RING zinc finger protein (RZF) mRNA, complete c
105F6	638	1209	AK026850	Hs.6906	0	1	FLJ23197 fis, clone REC00917 /cds=UNKNOW
178G10	5939	6469	AJ238403	Hs.6947	0	1	mRNA for huntingtin interacting protein 1 /cd
72A2	178	2992	AF001542	Hs.6975	0	9	AF001542 /clone=alpha_est218/52C1 /gb=
37F2	1757	2397	AK022568	Hs.7010	0	1	FLJ12506 fis, clone NT2RM2001700, weakly

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

598D3	1153	1299	NM_004637	Hs.7016	8.00E-56	1	RAB7, member RAS oncogene family (RAB7), mRNA
524C11	5542	5678	AB033034	Hs.7041	3.00E-72	1	mRNA for KIAA1208 protein, partial cds /cds=(2
109E10	452	1093	AF104921	Hs.7043	0	1	succinyl-CoA synthetase alpha subunit (SUCLA1
595F7	449	1150	NM_003849	Hs.7043	0	2	succinate-CoA ligase, GDP-forming, alpha sub
104H2	644	992	NM_020194	Hs.7045	1.00E-156	1	GL004 protein (GL004), mRNA /cds=(72,728) /gb
155C1	3322	3779	AK024478	Hs.7049	0	2	FLJ00071 protein, partial cds /cds=(3
473B1	3029	3439	AB051492	Hs.7076	1.00E-152	1	mRNA for KIAA1705 protein, partial cds /cds=(1
125E3	3612	3948	AL390127	Hs.7104	0	1	mRNA; cDNA DKFZp761P06121 (from clone DKFZp761
499B11	1451	1852	NM_021188	Hs.7137	0	2	clones 23667 and 23775 zinc finger protein (LOC
52B12	1850	2178	U90919	Hs.7137	1.00E-174	1	clones 23667 and 23775 zinc finger protein mRNA, compl
486A11	855	1186	NM_003904	Hs.7165	1.00E-132	1	zinc finger protein 259 (ZNF259), mRNA /cds=(2
460B6	2514	3182	NM_021931	Hs.7174	0	1	hypothetical protein FLJ22759 (FLJ22759), mR
592H8	3999	4524	AB051544	Hs.7187	0	2	mRNA for KIAA1757 protein, partial cds /cds=(3
180A10	102	468	AL117502	Hs.7200	1.00E-141	3	mRNA; cDNA DKFZp434D0935 (from clone DKFZp434
127A12	1503	2688	AL035661	Hs.7218	0	2	DNA sequence from clone RP4-568C11 on chromosome 20p1
592G9	12	263	NM_015953	Hs.7236	1.00E-138	2	CGI-25 protein (LOC51070), mRNA /cds=(44,949)
127E3	2624	4554	AB028980	Hs.7243	0	3	mRNA for KIAA1057 protein, partial cds /cds=(0
135F2	5029	5175	AB033050	Hs.7252	3.00E-78	1	mRNA for KIAA1224 protein, partial cds /cds=(0
57G1	2299	2723	NM_014319	Hs.7256	0	1	integral inner nuclear membrane protein (MAN1
122D11	2920	3123	AB014558	Hs.7278	5.00E-74	1	mRNA for KIAA0658 protein, partial cds /cds=(0
471H6	1	449	AV702692	Hs.7312	0	1	AV702692 cDNA, 5' end /clone=ADBBQC12 /clone_
104G12	4314	4797	AF084555	Hs.7351	0	2	okadaic acid-inducible and cAMP-regulated ph
590G7	771	1259	NM_005662	Hs.7381	0	5	voltage-dependent anion channel 3 (VDAC3), mR
159H2	355	1252	AL137423	Hs.7392	0	3	mRNA; cDNA DKFZp761E0323 (from clone DKFZp761E
161F3	1708	2371	NM_024045	Hs.7392	0	1	hypothetical protein MGC3199 (MGC3199), mRNA
195E1	1107	1362	NM_022736	Hs.7503	1.00E-129	1	hypothetical protein FLJ14153 (FLJ14153), mR
137F5	59	666	NM_018491	Hs.7535	0	2	COBW-like protein (LOC55871), mRNA /cds=(64,9
597E1	2302	2893	AF126028	Hs.7540	0	2	unknown mRNA /cds=(0,1261) /gb=AF126028 /gi=
473B6	3006	3302	AK025615	Hs.7567	1.00E-158	1	cDNA: FLJ21962 fis, clone HEP05564 /cds=UNKNOW
519H1	232	720	BG112505	Hs.7589	0	2	602282107F1 cDNA, 5' end /clone=IMAGE:4369729
73A9	106	3912	M20681	Hs.7594	0	8	glucose transporter-like protein-III (GLUT3), compl
51D3	106	3200	NM_006931	Hs.7594	0	2	solute carrier family 2 (facilitated glucose t
596E8	1512	1748	M94046	Hs.7647	1.00E-129	2	zinc finger protein (MAZ) mRNA /cds=UNKNOWN /gb=M9404
472A8	1575	1983	NM_004576	Hs.7688	0	1	protein phosphatase 2 (formerly 2A), regulator
191A10	386	889	NM_007278	Hs.7719	0	3	GABA(A) receptor-associated protein (GABARAP
459C4	5636	5897	AB002323	Hs.7720	2.00E-87	1	mRNA for KIAA0325 gene, partial cds /cds=(0,6265) /gb
99A12	606	1253	NM_018453	Hs.7731	0	1	uncharacterized bone marrow protein BM036 (BM
72G8	5806	6409	AB007938	Hs.7764	0	5	for KIAA0469 protein, complete cds /cds=(
45G2	6168	6404	NM_014851	Hs.7764	1.00E-132	1	KIAA0469 gene product (KIAA0469), mRNA /cds=(
172A4	371	588	NM_007273	Hs.7771	1.00E-107	1	B-cell associated protein (REA), mRNA /cds=(9
177B8	2055	2431	AK023166	Hs.7797	0	1	FLJ13104 fis, clone NT2RP3002343 /cds=(28
99B6	865	1244	NM_012461	Hs.7797	0	1	TERF1 (TRF1)-interacting nuclear factor 2 (T
160G8	727	860	U94855	Hs.7811	5.00E-66	1	translation initiation factor 3 47 kDa subunit
54G6	1	1007	AK001319	Hs.7837	1.00E-148	3	FLJ10457 fis, clone NT2RP1001424 /cds=UN
594A7	1295	1793	NM_013446	Hs.7838	0	4	makorin, ring finger protein, 1 (MKRN1), mRNA
188A12	1	2013	NM_017761	Hs.7862	0	3	hypothetical protein FLJ20312 (FLJ20312), mR
594A2	3060	3588	AK023813	Hs.7871	0	2	cDNA FLJ13751 fis, clone PLACE3000339, weakly
124C12	472	1251	NM_001550	Hs.7879	0	1	interferon-related developmental regulator
147A8	1381	1711	Y10313	Hs.7879	1.00E-134	1	for PC4 protein (IFRD1 gene) /cds=(219,158

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

74H3	4430	4978	AF302505	Hs.7886	0	2	pellino 1 (PELI1) mRNA, complete cds /cds=(4038
71G3	473	1112	NM_016224	Hs.7905	0	2	SH3 and PX domain-containing protein SH3PX1 (S
52C7	1637	2231	AB029551	Hs.7910	0	1	YEAF1 mRNA for YY1 and E4TF1 associated factor
177H5	5411	6045	AB002321	Hs.7911	0	1	KIAA0323 gene, partial cds /cds=(0,2175) /gb
114C8	1678	3078	NM_017657	Hs.7942	1.00E-149	2	hypothetical protein FLJ20080 (FLJ20080), mR
169D8	1453	2158	AK001437	Hs.7943	0	1	FLJ10575 fis, clone NT2RP2003295, highly
599G8	618	1204	NM_003796	Hs.7943	0	1	RPB5-mediating protein (RMP), mRNA /cds=(465,
127E11	107	796	NM_016099	Hs.7953	0	3	HSPC041 protein (LOC51125), mRNA /cds=(141,45
98D6	4769	6506	NM_001111	Hs.7957	0	20	adenosine deaminase, RNA-specific (ADAR), tr
37H10	2479	6594	X79448	Hs.7957	0	8	IFI-4 mRNA for type I protein /cds=(1165,3960) /g
178G4	4209	5132	AB028981	Hs.8021	0	4	mRNA for KIAA1058 protein, partial cds /cds=(0
118E9	630	1688	NM_006083	Hs.8024	0	2	IK cytokine, down-regulator of HLA II (IK), mRN
171A8	1658	1973	AK002026	Hs.8033	1.00E-151	1	FLJ11164 fis, clone PLACE1007226, weakly
103G5	1504	1977	NM_018346	Hs.8033	0	1	hypothetical protein FLJ11164 (FLJ11164), mR
179G7	2860	3032	AK022497	Hs.8068	6.00E-46	1	FLJ12435 fis, clone NT2RM1000059 /cds=(88
594A11	2327	2658	NM_018210	Hs.8083	1.00E-167	1	hypothetical protein FLJ10769 (FLJ10769), mR
103B5	1968	2448	AF267856	Hs.8084	0	1	HT033 mRNA, complete cds /cds=(203,931) /gb=A
98E4	1367	1808	AF113008	Hs.8102	0	7	clone FLB0708 mRNA sequence /cds=UNKNOWN /gb=
191H10	4581	5819	NM_018695	Hs.8117	0	3	erbb2-interacting protein ERBIN (LOC55914),
99F1	550	2672	AB014550	Hs.8118	0	4	mRNA for KIAA0650 protein, partial cds /cds=(0
165H11	488	663	NM_024408	Hs.8121	3.00E-93	1	Notch (Drosophila) homolog 2 (NOTCH2), mRNA /
515C7	2188	2514	AL050371	Hs.8128	1.00E-114	1	mRNA; cDNA DKFZp566G2246 (from clone DKFZp566G
166A12	234	1196	AF131856	Hs.8148	1.00E-155	2	clone 24856 mRNA sequence, complete cds /cds=(
520H8	512	712	NM_016275	Hs.8148	1.00E-110	1	selenoprotein T (LOC51714), mRNA /cds=(138,62
592D4	1	735	NM_014886	Hs.8170	1.00E-152	3	hypothetical protein (YR-29), mRNA /cds=(82,8
105F12	349	760	AK001665	Hs.8173	0	1	FLJ10803 fis, clone NT2RP4000833 /cds=(1
75A7	737	1458	AF000652	Hs.8180	0	1	syntenin (sycl) mRNA, complete cds /cds=(148,1
64H5	105	618	NM_005625	Hs.8180	0	3	syndecan binding protein (syntenin) (SDCBP),
61G9	3147	3660	AB018339	Hs.8182	0	2	for KIAA0796 protein, partial cds /cds=(0
39G2	255	1675	AF042284	Hs.8185	0	4	unknown mRNA /cds=(76,1428) /gb=AF042284 /gi
192G5	1054	1580	NM_021199	Hs.8185	0	8	CGI-44 protein; sulfide dehydrogenase like (y
109D3	1463	2503	AF269150	Hs.8203	0	2	transmembrane protein TM9SF3 (TM9SF3) mRNA, c
115H4	1251	3187	NM_020123	Hs.8203	0	12	endomembrane protein emp70 precursor isolog (
113F12	2349	3576	AL355476	Hs.8217	4.00E-35	2	DNA sequence from clone RP11-517O1 on chromosome X Co
125D5	582	1050	NM_005006	Hs.8248	0	1	NADH dehydrogenase (ubiquinone) Fe-S protein
460D3	4851	5043	AF035947	Hs.8257	7.00E-76	1	cytokine-inducible inhibitor of signalling t
111E7	729	3182	NM_013995	Hs.8262	0	2	lysosomal-associated membrane protein 2 (LAM
590F10	3012	4133	AK022790	Hs.8309	0	6	cDNA FLJ12728 fis, clone NT2RP2000040, highly
109B1	138	476	AW973507	Hs.8360	1.00E-161	1	EST385607 /gb=AW973507 /gi=8164686 /ug=
61A3	1137	1649	AB033017	Hs.8594	0	1	for KIAA1191 protein, partial cds /cds=(0
523E12	905	2998	NM_007271	Hs.8724	0	4	serine threonine protein kinase (NDR), mRNA /
590G2	3618	3932	NM_018031	Hs.8737	1.00E-166	3	WD repeat domain 6 (WDR6), mRNA /cds=(39,3404)
464C3	2299	2494	NM_018255	Hs.8739	1.00E-107	1	hypothetical protein FLJ10879 (FLJ10879), mR
128H8	1580	1711	NM_018450	Hs.8740	2.00E-64	1	uncharacterized bone marrow protein BM029 (BM
179D3	921	1457	AF083255	Hs.8765	0	1	RNA helicase-related protein complete c
195H11	1247	1481	NM_007269	Hs.8813	1.00E-100	1	syntaxin binding protein 3 (STXBP3), mRNA /cds
460F1	68	308	AA454036	Hs.8832	1.00E-105	1	zx48b04.r1 cDNA, 5' end /clone=IMAGE:795439 /
110E10	3672	5371	AB032252	Hs.8858	0	3	BAZ1A mRNA for bromodomain adjacent to zinc fi
113D1	4814	5890	NM_013448	Hs.8858	0	2	bromodomain adjacent to zinc finger domain, 1A
120H7	373	633	NM_017748	Hs.8928	1.00E-143	1	hypothetical protein FLJ20291 (FLJ20291), mR
470F10	1670	2260	NM_003917	Hs.8991	0	2	adaptor-related protein complex 1, gamma 2 su
72H11	1785	2418	M11717	Hs.8997	1.00E-147	23	heat shock protein (hsp 70) gene, complete cds /cds=(2

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

49H4	1769	2243	NM_005345	Hs.8997	1.00E-145	12	heat shock 70kD protein 1A (HSPA1A), mRNA /cds=
519E7	270	729	NM_003574	Hs.9006	0	1	VAMP (vesicle-associated membrane protein)-a
142E2	1265	1518	AK022215	Hs.9043	1.00E-107	1	FLJ12153 fis, clone MAMMA1000458 /cds=UNK
108B9	1160	1823	AJ002030	Hs.9071	0	1	for putative progesterone binding protein
47C7	452	795	AB011420	Hs.9075	0	1	for DRAK1, complete cds /cds=(117,1361) /
590A4	791	1377	NM_004760	Hs.9075	0	4	serine/threonine kinase 17a (apoptosis-induc
168D11	1000	1641	NM_017426	Hs.9082	0	1	nucleoporin p54 (NUP54), mRNA /cds=(25,1542)
63H9	799	1163	Y17829	Hs.9192	0	1	for Homer-related protein Syn47 /cds=(75,
167B11	1466	1863	NM_006251	Hs.9247	0	1	protein kinase, AMP-activated, alpha 1 cataly
196D5	1021	1492	AK024327	Hs.9343	0	1	cDNA FLJ14265 fis, clone PLACE1002256 /cds=UNK
192F3	245	790	NM_017983	Hs.9398	0	1	hypothetical protein FLJ10055 (FLJ10055), mR
121C3	3381	3567	AF217190	Hs.9414	3.00E-90	1	MLEL1 protein (MLEL1) mRNA, complete cds /cds=
196B6	959	1551	NM_003601	Hs.9456	0	1	SWI/SNF related, matrix associated, actin dep
331B5	2624	2950	AF027302	Hs.9573	1.00E-179	1	TNF-alpha stimulated ABC protein (ABC50) mRNA
592E11	1	479	NM_002520	Hs.9614	1.00E-139	7	nucleophosmin (nucleolar phosphoprotein B23
515D6	1739	2091	AB037796	Hs.9663	1.00E-160	1	mRNA for KIAA1375 protein, partial cds /cds=(0
124A5	1387	1762	NM_012068	Hs.9754	0	2	activating transcription factor 5 (ATF5), mRN
122A7	1484	1928	AB028963	Hs.9846	1.00E-154	1	mRNA for KIAA1040 protein, partial cds /cds=(0
591E2	1626	2194	AF123073	Hs.9851	0	5	C/EBP-induced protein mRNA, complete cds /cds
111G2	4208	5361	AB033076	Hs.9873	0	2	mRNA for KIAA1250 protein, partial cds /cds=(0
469D5	932	3551	AK022758	Hs.9908	1.00E-178	6	cDNA FLJ12696 fis, clone NT2RP1000513, highly
590D5	172	742	NM_001425	Hs.9999	2.00E-94	2	epithelial membrane protein 3 (EMP3), mRNA /c
112E7	1065	1753	NM_001814	Hs.10029	0	1	cathepsin C (CTSC), mRNA /cds=(33,1424) /gb=N
106C7	1066	1641	X87212	Hs.10029	0	1	cathepsin C /cds=(33,1424) /gb=X87212 /
127B1	1003	1429	NM_014959	Hs.10031	0	1	KIAA0955 protein (KIAA0955), mRNA /cds=(313,1
462E5	332	487	AW293461	Hs.10041	3.00E-46	1	UI-H-BI2-ahm-e-02-0-UI.s1 cDNA, 3' end /clon
190E3	101	356	NM_016551	Hs.10071	6.00E-98	1	seven transmembrane protein TM7SF3 (TM7SF3),
61B6	2571	2764	AL163249	Hs.10175	7.00E-94	1	chromosome 21 segment HS21C049 /cds=(128,2599
110F6	5310	5808	D87432	Hs.10315	0	1	KIAA0245 gene, complete cds /cds=(261,1808)
196E10	5312	5753	NM_003983	Hs.10315	0	1	solute carrier family 7 (cationic amino acid t
49D8	315	2207	AK024597	Hs.10362	0	3	cDNA: FLJ20944 fis, clone ADSE01780 /cds=UNKNO
129C7	1000	1364	AB018249	Hs.10458	0	1	CC chemokine LEC, complete cds /cds=(1
62F11	1239	2034	AL031685	Hs.10590	0	2	DNA sequence from clone RP5-963K23 on chromosome 20q1
460D5	86	815	AL357374	Hs.10600	0	4	DNA sequence from clone RP11-353C18 on chromosome 20
179C12	3765	4300	AK000005	Hs.10647	0	2	FLJ00005 protein, partial cds /cds=(0
482D12	1753	2359	NM_004848	Hs.10649	0	1	basement membrane-induced gene (ICB-1), mRNA
184F4	2686	3194	AL137721	Hs.10702	0	1	mRNA; cDNA DKFZp761H221 (from clone DKFZp761H2
186F10	2688	3084	NM_017601	Hs.10702	1.00E-137	2	hypothetical protein DKFZp761H221 (DKFZp761H
461E3	593	1110	NM_021821	Hs.10724	0	1	MDS023 protein (MDS023), mRNA /cds=(335,1018)
598D5	660	1191	NM_014306	Hs.10729	0	2	hypothetical protein (HSPC117), mRNA /cds=(75
125D9	104	397	NM_002495	Hs.10758	1.00E-165	1	NADH dehydrogenase (ubiquinone) Fe-S protein
36A7	172	1114	NM_006325	Hs.10842	0	11	RAN, member RAS oncogene familyRAN, member RAS
54H1	240	1467	NM_012257	Hs.10882	0	2	HMG-box containing protein 1 (HBP1), mRNA /cds
596B8	1186	1895	AK025212	Hs.10888	0	17	cDNA: FLJ21559 fis, clone COL06406 /cds=UNKNOW
458G7	989	1492	Z78330	Hs.10927	0	1	HSZ78330 cDNA /clone=2.49-(CEPH) /gb=Z78330
115D2	308	638	BF793378	Hs.10957	1.00E-102	1	602254823F1 cDNA, 5' end /clone=IMAGE:4347076
148H9	226	863	AF021819	Hs.10958	0	1	RNA-binding protein regulatory subunit mRNA,
173D5	356	816	NM_007262	Hs.10958	0	1	RNA-binding protein regulatory subunit (DJ-1
39B7	1553	2256	AF063605	Hs.11000	0	1	brain my047 protein mRNA, complete cds /cds=(8
592H5	1553	2257	NM_015344	Hs.11000	0	3	MY047 protein (MY047), mRNA /cds=(84,479) /gb

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

112G3	2591	3180	AB046813	Hs.11123	0	1	mRNA for KIAA1593 protein, partial cds /cds=(4
592E8	251	725	NM_014041	Hs.11125	0	2	HSPC033 protein (HSPC033), mRNA /cds=(168,443
477A2	1610	1697	NM_003100	Hs.11183	8.00E-43	2	sorting nexin 2 (SNX2), mRNA /cds=(29,1588) /g
41G4	6498	6751	AB014522	Hs.11238	1.00E-142	1	for KIAA0622 protein, partial cds /cds=(0
519A3	759	987	NM_018371	Hs.11260	1.00E-127	1	hypothetical protein FLJ11264 (FLJ11264), mR
175B4	404	688	BE788546	Hs.11355	4.00E-75	1	601476186F1 cDNA, 5' end /clone=IMAGE:3878948
114F11	245	401	BF665055	Hs.11356	4.00E-55	1	602119656F1 cDNA, 5' end /clone=IMAGE:4276860
40D2	96	824	U59808	Hs.11383	0	1	monocyte chemotactic protein-4 precursor (MCP-4) mR
109C3	767	2345	M74002	Hs.11482	0	2	arginine-rich nuclear protein mRNA, complete cds /cds
117G9	408	2345	NM_004768	Hs.11482	0	8	splicing factor, arginine/serine-rich 11 (SF
458G6	2053	2164	AK022628	Hs.11556	1.00E-54	1	cDNA FLJ12566 fis, clone NT2RM4000852 /cds=UNK
181E7	644	1004	AK021632	Hs.11571	1.00E-167	1	cDNA FLJ11570 fis, clone HEMBA1003309 /cds=UNK
458B3	85	522	R12665	Hs.11594	1.00E-137	1	yf40a04.s1 cDNA, 3' end /clone=IMAGE:129294 /
146B6	498	677	BE794595	Hs.11607	5.00E-82	1	601590368F1 5' end /clone=IMAGE:3944489
516F12	388	711	BG288429	Hs.11637	1.00E-132	1	602388093F1 cDNA, 5' end /clone=IMAGE:4517086
60B1	1291	1882	NM_005121	Hs.11861	0	1	thyroid hormone receptor-associated protein,
44C6	2613	2834	NM_000859	Hs.11899	9.00E-72	1	3-hydroxy-3-methylglutaryl-Coenzyme A reduc
39F10	1	221	BF668230	Hs.12035	1.00E-120	2	602122419F1 cDNA, 5' end /clone=IMAGE:4279300
596D8	234	849	U72514	Hs.12045	0	2	C2f mRNA, complete cds
481E7	1902	2190	AB028986	Hs.12064	1.00E-151	1	mRNA for KIAA1063 protein, partial cds /cds=(0
465D9	2529	2699	NM_004003	Hs.12068	8.00E-91	1	carnitine acetyltransferase (CRAT), nuclear
116H8	283	738	NM_003321	Hs.12084	0	1	Tu translation elongation factor, mitochondri
44A4	319	836	S75463	Hs.12084	0	1	P43=mitochondrial elongation factor homolog [human, live
114F7	4254	4495	AL137753	Hs.12144	1.00E-115	1	mRNA; cDNA DKFZp434K1412 (from clone DKFZp434K
123F12	1	219	NM_021203	Hs.12152	1.00E-114	1	APMCF1 protein (APMCF1), mRNA /cds=(82,225) /
519H7	166	753	AK025775	Hs.12245	0	1	cDNA: FLJ22122 fis, clone HEP19214 /cds=UNKNOW
70E3	953	4720	AB014530	Hs.12259	0	3	for KIAA0630 protein, partial cds /cds=(0
107H1	680	1078	AK024756	Hs.12293	0	1	FLJ21103 fis, clone CAS04883 /cds=(107,1
71E5	4750	5283	NM_003170	Hs.12303	0	1	suppressor of Ty (S.cerevisiae) 6 homolog (SUP
106F3	977	1490	AL050272	Hs.12305	0	1	cDNA DKFZp566B183 (from clone DKFZp566B1
481F4	1859	2403	NM_015509	Hs.12305	0	1	DKFZP566B183 protein (DKFZP566B183), mRNA /c
114D3	1271	1520	AF038202	Hs.12311	1.00E-118	1	clone 23570 mRNA sequence /cds=UNKNOWN /gb=AF0
463B9	1006	1224	AK021670	Hs.12315	1.00E-121	1	cDNA FLJ11608 fis, clone HEMBA1003976 /cds=(56
167A8	71	723	BG034192	Hs.12396	0	2	602302446F1 cDNA, 5' end /clone=IMAGE:4403866
460E9	3808	4166	D83776	Hs.12413	1.00E-176	1	mRNA for KIAA0191 gene, partial cds /cds=(0,4552) /gb
157E1	1887	3154	NM_020403	Hs.12450	0	3	cadherin superfamily protein VR4-11 (LOC57123
69F11	2715	3447	AK001676	Hs.12457	0	1	FLJ10814 fis, clone NT2RP4000984 /cds=(92
118B8	5781	6374	AB032973	Hs.12461	0	1	mRNA for KIAA1147 protein, partial cds /cds=(0
193G12	2069	2368	NM_005993	Hs.12570	1.00E-169	1	tubulin-specific chaperone d (TBCD), mRNA /cd
459D11	2828	3122	NM_021151	Hs.12743	1.00E-147	1	carnitine octanoyltransferase (COT), mRNA /c
196H4	1	5439	AB046785	Hs.12772	0	2	mRNA for KIAA1565 protein, partial cds /cds=(0
56G11	458	1088	AL080156	Hs.12813	0	1	cDNA DKFZp434J214 (from clone DKFZp434J2
476E6	1221	1638	NM_006590	Hs.12820	0	1	SnRNP assembly defective 1 homolog (SAD1), mRN
109E7	1	180	AF208855	Hs.12830	3.00E-79	1	BM-013 mRNA, complete cds /cds=(67,459) /gb=A

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

458A2	1818	2276	AK026747	Hs.12969	0	1	cDNA: FLJ23094 fis, clone LNG07379, highly sim
466D10	1469	1745	AK001822	Hs.12999	9.00E-39	1	cDNA FLJ10960 fis, clone PLACE1000564 /cds=UNK
187A11	1866	2555	NM_003330	Hs.13046	0	2	thioredoxin reductase 1 (TXNRD1), mRNA /cds=(
60D9	1757	3508	X91247	Hs.13046	0	3	thioredoxin reductase /cds=(439,1932)
75D7	2071	2550	AF055581	Hs.13131	0	1	adaptor protein Lnk mRNA, complete cds /cds=(3
196C2	190	845	AK026239	Hs.13179	0	2	cDNA: FLJ22586 fis, clone HSI02774 /cds=UNKNOW
480G6	11	380	AL570416	Hs.13256	1.00E-161	1	AL570416 cDNA /clone=CS0DI020YK05-(3-prime)
196H3	2814	3382	AB020663	Hs.13264	0	1	mRNA for KIAA0856 protein, partial cds /cds=(0
460H3	127	431	BF029796	Hs.13268	1.00E-151	1	601556721F1 cDNA, 5' end /clone=IMAGE:3826637
170B2	1487	1635	AB011164	Hs.13273	1.00E-69	1	for KIAA0592 protein, partial cds /cds=(0,
115E6	2153	2376	AK025707	Hs.13277	1.00E-124	1	cDNA: FLJ22054 fis, clone HEP09634 /cds=(144,9
110F10	119	648	BE537908	Hs.13328	0	1	601067373F1 cDNA, 5' end /clone=IMAGE:3453594
36C2	427	4137	AF054284	Hs.13453	0	5	spliceosomal protein SAP 155 mRNA, complete cd
594C3	5	4229	NM_012433	Hs.13453	0	10	splicing factor 3b, subunit 1, 155kD (SF3B1), m
110C6	4	1853	AF131753	Hs.13472	0	5	clone 24859 mRNA sequence /cds=UNKNOWN /gb=AF
173B6	1156	1672	NM_013236	Hs.13493	0	1	like mouse brain protein E46 (E46L), mRNA /cds=
462C4	794	1093	BC001909	Hs.13580	1.00E-115	1	clone IMAGE:3537447, mRNA, partial cds /cds=
597H11	412	936	NM_014174	Hs.13645	0	1	HSPC144 protein (HSPC144), mRNA /cds=(446,112
107F8	429	821	AK025767	Hs.13755	0	1	FLJ22114 fis, clone HEP18441 /cds=UNKNOW
102D12	3153	4764	AF000993	Hs.13980	0	2	ubiquitous TPR motif, X isoform (UTX) mRNA, alt
515G12	1710	2120	AK025425	Hs.14040	0	2	cDNA: FLJ21772 fis, clone COLF7808 /cds=UNKNOW
480H5	1945	2259	AK024228	Hs.14070	1.00E-119	1	cDNA FLJ14166 fis, clone NT2RP1000796 /cds=(20
61D1	73	499	NM_014245	Hs.14084	0	1	ring finger protein 7 (RNF7), mRNA /cds=(53,394
122E4	2162	2685	NM_014454	Hs.14125	0	1	p53 regulated PA26 nuclear protein (PA26), mRN
123D9	22	722	NM_001161	Hs.14142	0	1	nudix (nucleoside diphosphate linked moiety
460F11	1084	1322	NM_017827	Hs.14220	4.00E-74	1	hypothetical protein FLJ20450 (FLJ20450), mR
458D2	127	536	NM_018648	Hs.14317	0	1	nucleolar protein family A, member 3 (H/ACA sm
167G1	30	198	AK022939	Hs.14347	3.00E-91	1	cDNA FLJ12877 fis, clone NT2RP2003825 /cds=(3
117H10	975	1721	NM_003022	Hs.14368	0	1	SH3 domain binding glutamic acid-rich protein
591B12	1082	1801	NM_001614	Hs.14376	0	9	actin, gamma 1 (ACTG1), mRNA /cds=(74,1201) /g
179H3	1160	1791	X04098	Hs.14376	1.00E-178	5	cytoskeletal gamma-actin /cds=(73,1200) /g
116D9	5818	6073	NM_012199	Hs.14520	5.00E-84	1	eukaryotic translation initiation factor 2C,
64D11	1901	2506	NM_003592	Hs.14541	0	1	cullin 1 (CUL1), mRNA /cds=(124,2382) /gb=NM_0
516F4	750	1331	AK025166	Hs.14555	0	1	cDNA: FLJ21513 fis, clone COL05778 /cds=UNKNOW
459G5	1	260	AK025269	Hs.14562	5.00E-88	1	cDNA: FLJ21616 fis, clone COL07477 /cds=(119,1
521B7	7	1825	NM_005335	Hs.14601	0	8	hematopoietic cell-specific Lyn substrate 1
110D7	7	1295	X16663	Hs.14601	0	3	HS1 gene for heamatopoietic lineage cell specific pro
114D11	1460	1559	NM_003584	Hs.14611	1.00E-45	1	dual specificity phosphatase 11 (RNA/RNP comp
589A3	1665	2197	NM_016293	Hs.14770	0	2	bridging integrator 2 (BIN2), mRNA /cds=(38,17
104C8	2113	2380	AB031050	Hs.14805	1.00E-135	2	for organic anion transporter OATP-D, com
481D10	2466	2694	NM_013272	Hs.14805	1.00E-68	1	solute carrier family 21 (organic anion transp
125B2	2704	3183	NM_001455	Hs.14845	0	1	forkhead box O3A (FOXO3A), mRNA /cds=(924,2945
500D7	2174	2379	AL050021	Hs.14846	1.00E-100	1	mRNA; cDNA DKFZp564D016 (from clone DKFZp564D0
123B5	1793	2195	NM_016598	Hs.14896	0	1	DHHC1 protein (LOC51304), mRNA /cds=(214,1197
499E2	1266	1549	AB020644	Hs.14945	1.00E-155	3	mRNA for KIAA0837 protein, partial cds /cds=(0
123H6	2980	3652	NM_007192	Hs.14963	0	3	chromatin-specific transcription elongation
61G10	264	528	D13627	Hs.15071	1.00E-144	1	KIAA0002 gene, complete cds /cds=(28,1674) /

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

460D10	2162	4305	NM_014837	Hs.15087	0	4	KIAA0250 gene product (KIAA0250), mRNA /cds=(
176E12	9289	9739	NM_022473	Hs.15220	0	1	zinc finger protein 106 (ZFP106), mRNA /cds=(3
487E11	1561	1989	NM_006170	Hs.15243	0	1	nucleolar protein 1 (120kD) (NOL1), mRNA /cds=
75E11	1628	2201	AF127139	Hs.15259	0	20	Bcl-2-binding protein BIS (BIS) mRNA, complete
71H9	1656	2532	NM_004281	Hs.15259	0	12	BCL2-associated athanogene 3 (BAG3), mRNA /cd
484G9	465	1006	NM_005826	Hs.15265	0	1	heterogeneous nuclear ribonucleoprotein R (
480H8	2013	2635	AB037828	Hs.15370	0	1	mRNA for KIAA1407 protein, partial cds /cds=(0
587G9	2436	2769	AK024088	Hs.15423	1.00E-167	1	cDNA FLJ14026 fis, clone HEMBA1003679, weakly
483D6	5239	5810	NM_004774	Hs.15589	0	1	PPAR binding protein (PPARBP), mRNA /cds=(235,
514A7	673	942	NM_006833	Hs.15591	1.00E-151	1	COP9 subunit 6 (MOV34 homolog, 34 kD) (MOV34-34
125A2	522	746	NM_024348	Hs.15961	1.00E-112	1	dynactin 3 (p22) (DCTN3), transcript variant
591A5	295	704	NM_005005	Hs.15977	0	3	NADH dehydrogenase (ubiquinone) 1 beta subcom
39H12	1641	1993	X74262	Hs.16003	1.00E-180	1	RbAp48 mRNA encoding retinoblastoma binding prot
113A9	1328	1891	NM_016334	Hs.16085	0	1	putative G-protein coupled receptor (SH120),
45C2	765	1674	NM_006461	Hs.16244	0	2	mitotic spindle coiled-coil related protein (
494H10	113	2576	NM_016312	Hs.16420	0	3	Npw38-binding protein NpwBP (LOC51729), mRNA
40D8	52	246	Y13710	Hs.16530	1.00E-107	1	for alternative activated macrophage spe
597E7	244	524	AL523085	Hs.16648	1.00E-147	1	AL523085 cDNA /clone=CS0DC001YF21-(5-prime)
458D11	232	319	AY007106	Hs.16773	1.00E-42	1	clone TCCCIA00427 mRNA sequence /cds=UNKNOWN
70F2	824	991	AL021786	Hs.17109	2.00E-90	2	DNA sequence from PAC 696H22 on chromosome Xq21.1-21.2
167C5	5768	5905	D86964	Hs.17211	3.00E-62	1	mRNA for KIAA0209 gene, partial cds /cds=(0,5530) /gb
460H2	3424	3624	AL162070	Hs.17377	1.00E-103	1	mRNA; cDNA DKFZp762H186 (from clone DKFZp762H1
70G11	1384	1885	AK023680	Hs.17448	0	2	FLJ13618 fis, clone PLACE1010925 /cds=UNK
129C11	2458	3044	U47924	Hs.17483	0	2	chromosome 12p13 sequence /cds=(194,1570) /gb=U4792
467H3	4713	4908	NM_014521	Hs.17667	1.00E-61	1	SH3-domain binding protein 4 (SH3BP4), mRNA /
71A11	100	370	BG035218	Hs.17719	1.00E-142	1	602324727F1 cDNA, 5' end /clone=IMAGE:4412910
598C7	513	902	NM_021622	Hs.17757	1.00E-178	1	pleckstrin homology domain-containing, fami
595A7	3296	5680	AB046774	Hs.17767	0	5	mRNA for KIAA1554 protein, partial cds /cds=(0
58D12	5225	5857	AB007861	Hs.17803	0	1	KIAA0401 mRNA, partial cds /cds=(0,1036) /gb=
524G8	357	809	NM_014350	Hs.17839	0	1	TNF-induced protein (GG2-1), mRNA /cds=(197,7
521B10	1008	1476	NM_002707	Hs.17883	0	2	protein phosphatase 1G (formerly 2C), magnesiu
69B12	1014	1490	Y13936	Hs.17883	0	1	for protein phosphatase 2C gamma /cds=(24,
178E6	1903	4365	NM_014827	Hs.17969	0	3	KIAA0663 gene product (KIAA0663), mRNA /cds=(
173H3	481	2362	AK001630	Hs.18063	0	4	cDNA FLJ10768 fis, clone NT2RP4000150 /cds=UN
113A8	1285	1393	NM_005606	Hs.18069	5.00E-48	1	protease, cysteine, 1 (legumain) (PRSC1), mRN
118H9	3709	3950	AB020677	Hs.18166	1.00E-125	1	mRNA for KIAA0870 protein, partial cds /cds=(0
513H7	2204	2757	NM_005839	Hs.18192	1.00E-112	3	Ser/Arg-related nuclear matrix protein (plen
523G9	507	768	AB044661	Hs.18259	1.00E-147	1	XAB1 mRNA for XPA binding protein 1, complete c
105B9	695	1115	AJ010842	Hs.18259	0	1	for putative ATP(GTP)-binding protein, p
589D12	335	715	NM_016565	Hs.18552	0	2	E2IG2 protein (LOC51287), mRNA /cds=(131,421)
170C8	414	737	AF072860	Hs.18571	0	2	protein activator of the interferon-induced p
189A12	414	736	NM_003690	Hs.18571	0	1	protein kinase, interferon-inducible double
134B9	2751	3057	AB046808	Hs.18587	1.00E-165	1	mRNA for KIAA1588 protein, partial cds /cds=(2
519G5	1291	1581	NM_012332	Hs.18625	1.00E-157	2	Mitochondrial Acyl-CoA Thioesterase (MT-ACT4
526H2	827	1205	NM_004208	Hs.18720	0	1	programmed cell death 8 (apoptosis-inducing f

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

462F12	409	556	NM_017899	Hs.18791	2.00E-78	1	hypothetical protein FLJ20607 (FLJ20607), mR
138B2	388	995	AF003938	Hs.18792	0	1	thioredoxin-like protein complete cds
36G12	935	1272	AJ250014	Hs.18827	0	2	for Familial Cylindromatosis cyld gene /
194D3	924	2123	NM_018253	Hs.18851	0	2	hypothetical protein FLJ10875 (FLJ10875), mR
523E1	3653	4056	NM_012290	Hs.18895	0	1	tousled-like kinase 1 (TLK1), mRNA /cds=(212,2
587G5	1	350	NM_016302	Hs.18925	1.00E-166	1	protein x 0001 (LOC51185), mRNA /cds=(33,1043)
595C10	161	1281	AC006042	Hs.18987	0	4	BAC clone RP11-505D17 from 7p22-p21 /cds=(0,12
125G10	54	752	NM_002492	Hs.19236	0	3	NADH dehydrogenase (ubiquinone) 1 beta subcom
478G7	1	193	NM_021603	Hs.19520	9.00E-51	1	FXYD domain-containing ion transport regulat
595F11	3623	3736	AB051481	Hs.19597	3.00E-49	1	mRNA for KIAA1694 protein, partial cds /cds=(0
177C6	284	671	AF161339	Hs.19807	0	2	HSPC076 mRNA, partial cds /cds=(0,301) /gb=AF
37E12	3485	3919	AB018298	Hs.19822	0	1	for KIAA0755 protein, complete cds /cds=(
64G8	962	1311	NM_001902	Hs.19904	0	1	cystathionase (cystathionine gamma-lyase) (
499D5	2829	3183	AB011169	Hs.20141	0	1	mRNA for KIAA0597 protein, partial cds /cds=(0,
40D11	62	684	NM_004166	Hs.20144	0	1	small inducible cytokine subfamily A (Cys-Cys
66C10	1240	2240	U76248	Hs.20191	0	12	hSIAH2 mRNA, complete cds /cds=(526,1500) /gb=U76248
586B12	1686	4288	AB040922	Hs.20237	0	2	mRNA for KIAA1489 protein, partial cds /cds=(1
173G8	2578	3197	AL096776	Hs.20252	0	1	DNA sequence from clone RP4-646B12 on chromosome 1q42
98C6	3303	4699	AB051487	Hs.20281	0	6	mRNA for KIAA1700 protein, partial cds /cds=(1
107H11	781	1380	AK022103	Hs.20281	0	1	FLJ12041 fis, clone HEMBB1001945 /cds=UNK
121B8	778	1264	NM_001548	Hs.20315	0	1	interferon-induced protein with tetratricope
110C4	1050	1431	AF244137	Hs.20597	0	1	hepatocellular carcinoma-associated antigen
99H6	899	1412	NM_014315	Hs.20597	0	2	host cell factor homolog (LCP), mRNA /cds=(316,
152B12	69	424	AK025446	Hs.20760	0	1	FLJ21793 fis, clone HEP00466 /cds=UNKNOW
459A8	1858	2143	AL021366	Hs.20830	1.00E-155	1	DNA sequence from cosmid ICK0721Q on chromosome
587A11	720	1080	AL137576	Hs.21015	0	1	mRNA; cDNA DKFZp564L0864 (from clone DKFZp564L
191E12	1688	2235	AK025019	Hs.21056	0	2	cDNA: FLJ21366 fis, clone COL03012, highly sim
52G3	225	1652	NM_005880	Hs.21189	0	6	HIRA interacting protein 4 (dnaJ-like) (HIRIP
181B7	3176	3316	AB018325	Hs.21264	3.00E-72	1	mRNA for KIAA0782 protein, partial cds /cds=(0
45E11	1378	1518	NM_003115	Hs.21293	1.00E-72	1	UDP-N-acteylglucosamine pyrophosphorylase
109G1	2989	3487	AB032948	Hs.21356	0	1	for KIAA1122 protein, partial cds /cds=(0
116D4	5522	5741	NM_016936	Hs.21479	1.00E-107	1	ubinuclein 1 (UBN1), mRNA /cds=(114,3518) /gb
37G10	294	3960	M97935	Hs.21486	0	4	transcription factor ISGF-3 mRNA, complete cd
599E8	329	3568	NM_007315	Hs.21486	0	6	signal transducer and activator of transcripti
592D10	2223	3204	NM_002709	Hs.21537	0	3	protein phosphatase 1, catalytic subunit, bet
68A7	1327	1612	AB028958	Hs.21542	1.00E-161	1	for KIAA1035 protein, partial cds /cds=(0
72B3	2519	2862	L03426	Hs.21595	1.00E-179	1	XE7 mRNA, complete alternate coding regions /cds=(166
592E6	2520	2854	NM_005088	Hs.21595	1.00E-161	1	DNA segment on chromosome X and (unique) 155 ex
589G6	190	522	AL573787	Hs.21732	1.00E-141	1	AL573787 cDNA /clone=CS0DI055YM17-(3-prime)
593H1	452	899	NM_005875	Hs.21756	0	2	translation factor sui1 homolog (GC20), mRNA
59B8	2893	3273	NM_012406	Hs.21807	0	1	PR domain containing 4 (PRDM4), mRNA /cds=(122,
196A9	12	543	AL562895	Hs.21812	0	1	AL562895 cDNA /clone=CS0DC021YO20-(3-prime)
67D8	62	631	AW512498	Hs.21879	1.00E-150	3	xx75e03.x1 cDNA, 3' end /clone=IMAGE:2849500
477B6	1969	2520	D84454	Hs.21899	0	1	mRNA for UDP-galactose translocator, complete cds /c
515D1	2232	2647	NM_007067	Hs.21907	0	2	histone acetyltransferase (HBOA), mRNA /cds=
100F8	1082	1508	AK022554	Hs.21938	0	1	FLJ12492 fis, clone NT2RM2001632, weakly
470E4	1135	1244	NM_020239	Hs.22065	4.00E-45	2	small protein effector 1 of Cdc42 (SPEC1), mRNA
68G4	1391	2013	AK022057	Hs.22265	0	2	FLJ11995 fis, clone HEMBB1001443, highly
193H6	922	1328	NM_022494	Hs.22353	1.00E-178	1	hypothetical protein FLJ21952 (FLJ21952), mR
151D2	1492	1694	AL049951	Hs.22370	4.00E-88	1	cDNA DKFZp564O0122 (from clone DKFZp564O

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

497E8	1581	4794	D83781	Hs.22559	0	3	mRNA for KIAA0197 gene, partial cds /cds=(0,3945) /gb
182D10	999	1830	AL117513	Hs.22583	0	5	mRNA; cDNA DKFZp434K2235 (from clone DKFZp434K
75B5	1775	2380	AF006513	Hs.22670	0	1	CHD1 mRNA, complete cds /cds=(163,5292) /gb=A
126H8	1776	2377	NM_001270	Hs.22670	0	1	chromodomain helicase DNA binding protein 1 (
73D5	1599	1696	AK025485	Hs.22678	2.00E-42	1	FLJ21832 fis, clone HEP01571 /cds=(32,15
481D11	128	562	BF968270	Hs.22790	1.00E-172	1	602269653F1 cDNA, 5' end /clone=IMAGE:4357740
74E4	724	1195	NM_012124	Hs.22857	0	1	chord domain-containing protein 1 (CHP1), mRN
459C6	813	1472	NM_012244	Hs.22891	0	1	solute carrier family 7 (cationic amino acid t
462G7	2972	3144	AB037784	Hs.22941	2.00E-93	1	mRNA for KIAA1363 protein, partial cds /cds=(0
70F12	37	846	AB020623	Hs.22960	0	3	DAM1 mRNA, complete cds /cds=(48,725) /gb=AB0
585H10	91	748	NM_005872	Hs.22960	0	1	breast carcinoma amplified sequence 2 (BCAS2)
142C8	1359	1597	AK024023	Hs.23170	1.00E-103	1	FLJ13961 fis, clone Y79AA1001236, highly
164F2	1220	1474	NM_012280	Hs.23170	1.00E-135	1	homolog of yeast SPB1 (JM23), mRNA /cds=(300,12
127F11	682	806	AL046016	Hs.23247	2.00E-58	1	DKFZp434P246_r1 cDNA, 5' end /clone=DKFZp434P
98G7	760	1368	NM_022496	Hs.23259	0	1	hypothetical protein FLJ13433 (FLJ13433), mR
470C9	2	538	AL574514	Hs.23294	0	2	AL574514 cDNA /clone=CS0DI056YA07-(3-prime)
458F12	4293	4917	AB002365	Hs.23311	0	1	mRNA for KIAA0367 gene, partial cds /cds=(0,2150) /gb
57D8	460	566	BF439063	Hs.23349	3.00E-54	1	nab70e03.x1 cDNA /clone=IMAGE /gb=BF439063 /
599G12	352	983	NM_014814	Hs.23488	0	2	KIAA0107 gene product (KIAA0107), mRNA /cds=(
112B3	2400	2715	NM_014887	Hs.23518	1.00E-172	1	hypothetical protein from BCRA2 region (CG005
167C10	1771	2107	NM_004380	Hs.23598	1.00E-175	1	CREB binding protein (Rubinstein-Taybi syndr
196G9	114	307	BF970427	Hs.23703	1.00E-101	1	602272760F1 cDNA, 5' end /clone=IMAGE:4360767
184B3	2488	2882	AK026983	Hs.23803	0	1	FLJ23330 fis, clone HEP12654 /cds=(69,13
480H4	4871	5467	AB023227	Hs.23860	0	1	mRNA for KIAA1010 protein, partial cds /cds=(0
479C12	4	190	NM_005556	Hs.23881	4.00E-91	1	keratin 7 (KRT7), mRNA /cds=(56,1465) /gb=NM_
36E7	742	1126	AL360135	Hs.23964	0	1	full length insert cDNA clone EUROIMAGE 12
598B5	544	1271	NM_005870	Hs.23964	0	12	sin3-associated polypeptide, 18kD (SAP18), m
462D8	1205	1653	NM_004790	Hs.23965	0	1	solute carrier family 22 (organic anion transp
479A5	1817	2164	NM_002967	Hs.23978	0	1	scaffold attachment factor B (SAFB), mRNA /cds
188E2	1762	2160	NM_014950	Hs.24083	0	1	KIAA0997 protein (KIAA0997), mRNA /cds=(262,2
67D2	1304	1856	AK024240	Hs.24115	0	2	FLJ14178 fis, clone NT2RP2003339 /cds=UNK
177D8	4674	5185	AF251039	Hs.24125	0	1	putative zinc finger protein mRNA, complete cd
190E1	5222	5394	NM_016604	Hs.24125	8.00E-73	1	putative zinc finger protein (LOC51780), mRNA
192A5	1517	1985	NM_003387	Hs.24143	1.00E-135	2	Wiskott-Aldrich syndrome protein interacting
170A4	1666	3280	X86019	Hs.24143	4.00E-23	1	PRPL-2 protein /cds=(204,1688) /gb=X860
480B6	1517	1937	NM_012155	Hs.24178	1.00E-133	1	microtubule-associated protein like echinode
143H11	177	656	BE877357	Hs.24181	0	2	601485590F1 cDNA, 5' end /clone=IMAGE:3887951
473D10	146	491	AW960486	Hs.24252	0	1	EST372557 cDNA /gb=AW960486 /gi=8150170 /ug=
98H1	23	562	NM_003945	Hs.24322	0	1	ATPase, H+ transporting, lysosomal (vacuolar
169G2	391	638	BE612847	Hs.24349	4.00E-75	2	601452239F1 5' end /clone=IMAGE:3856304
479B12	1132	1599	AY007126	Hs.24435	0	1	clone CDABP0028 mRNA sequence /cds=UNKNOWN /g
480H9	4716	5012	NM_006048	Hs.24594	1.00E-145	1	ubiquitination factor E4B (homologous to yeas
110B10	520	1171	AL163206	Hs.24633	0	1	chromosome 21 segment HS21C006 /cds=(82,1203)
99A3	519	1000	NM_022136	Hs.24633	0	2	SAM domain, SH3 domain and nuclear localisation
109G7	2024	2350	AB037797	Hs.24684	1.00E-141	1	for KIAA1376 protein, partial cds /cds=(1
61B7	485	1656	AK024029	Hs.24719	0	4	FLJ13967 fis, clone Y79AA1001402, weakly
166C11	1216	1509	AF006516	Hs.24752	1.00E-165	1	eps8 binding protein e3B1 mRNA, complete cds /

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

464D12	166	764	NM_002882	Hs.24763	0	1	RAN binding protein 1 (RANBP1), mRNA /cds=(149
98C12	6523	8023	AB051512	Hs.25127	0	3	mRNA for KIAA1725 protein, partial cds /cds=(0
63F7	2164	2802	AL133611	Hs.25362	0	1	cDNA DKFZp434O1317 (from clone DKFZp434O
41D11	45	463	X53795	Hs.25409	0	1	R2 mRNA for an inducible membrane protein /cds=(156,95
62G6	1452	1827	V01512	Hs.25647	0	3	cellular oncogene c-fos (complete sequence) /cds=(15
593D12	1135	2111	NM_015832	Hs.25674	0	8	methyl-CpG binding domain protein 2 (MBD2), tr
172G9	2014	2371	NM_021211	Hs.25726	0	1	transposon-derived Buster1 transposase-like
106D6	432	1878	AF058696	Hs.25812	0	2	cell cycle regulatory protein p95 (NBS1) mRNA,
98A4	533	3758	NM_002485	Hs.25812	0	2	Nijmegen breakage syndrome 1 (nibrin) (NBS1),
477H5	6320	6599	NM_004638	Hs.25911	1.00E-111	3	HLA-B associated transcript-2 (D6S51E), mRNA
71F11	2070	2931	NM_019555	Hs.25951	0	3	Rho guanine nucleotide exchange factor (GEF)
164B9	2163	2502	AK023999	Hs.26039	1.00E-159	1	cDNA FLJ13937 fis, clone Y79AA1000805 /cds=UNK
100A3	2043	2620	M34668	Hs.26045	0	1	protein tyrosine phosphatase (PTPase-alpha) mRNA /c
123A5	2046	2638	NM_002836	Hs.26045	0	1	protein tyrosine phosphatase, receptor type,
466E5	7817	8241	NM_014112	Hs.26102	0	2	trichorhinophalangeal syndrome I gene (TRPS1)
588A1	361	857	AF070582	Hs.26118	0	1	clone 24766 mRNA sequence /cds=UNKNOWN /gb=AF
526H12	176	1809	NM_018384	Hs.26194	0	5	hypothetical protein FLJ11296 (FLJ11296), mR
149G7	96	1123	AK027016	Hs.26198	0	3	FLJ23363 fis, clone HEP15507 /cds=(206,1
122A4	1196	1332	AL050166	Hs.26295	3.00E-72	1	mRNA; cDNA DKFZp586D1122 (from clone DKFZp586D
122D5	1936	2435	AB029006	Hs.26334	0	1	mRNA for KIAA1083 protein, complete cds /cds=(
137G5	137	452	AK025778	Hs.26367	1.00E-145	1	FLJ22125 fis, clone HEP19410 /cds=(119,5
595D2	1	372	NM_022488	Hs.26367	3.00E-89	3	PC3-96 protein (PC3-96), mRNA /cds=(119,586)
64D12	1024	1135	NM_017746	Hs.26369	2.00E-57	1	hypothetical protein FLJ20287 (FLJ20287), mR
39E4	2132	2750	AK000367	Hs.26434	0	1	FLJ20360 fis, clone HEP16677 /cds=(79,230
473C10	4318	4623	AF051782	Hs.26584	1.00E-154	1	diaphanous 1 (HDIA1) mRNA, complete cds /cds=(
590C4	1740	2198	AL050205	Hs.26613	0	1	mRNA; cDNA DKFZp586F1323 (from clone DKFZp586F
523F3	454	792	AC002073	Hs.26670	1.00E-164	1	PAC clone RP3-515N1 from 22q11.2-q22 /cds=(0,791) /g
587E11	1226	1876	NM_004779	Hs.26703	0	2	CCR4-NOT transcription complex, subunit 8 (C
110G4	191	685	BE868389	Hs.26731	0	1	601444360F1 cDNA, 5' end /clone=IMAGE:3848487
110E11	1001	3955	AL117448	Hs.26797	0	2	cDNA DKFZp586B1417 (from clone DKFZp586B
152A8	12	112	AI760224	Hs.26873	2.00E-48	1	wh62g06.x1 cDNA, 3' end /clone=IMAGE:2385370
467G11	528	858	NM_016106	Hs.27023	1.00E-174	1	vesicle transport-related protein (KIAA0917)
465E11	634	1065	AL136656	Hs.27181	3.00E-83	1	mRNA; cDNA DKFZp564C1664 (from clone DKFZp564C
58E11	1	551	AJ238243	Hs.27182	0	1	mRNA for phospholipase A2 activating protein
590H2	398	1016	NM_014412	Hs.27258	0	1	calcyclin binding protein (CACYBP), mRNA /cds
179E9	1039	1905	AK025586	Hs.27268	0	4	FLJ21933 fis, clone HEP04337 /cds=UNKNOW
459D7	1293	1936	AL050061	Hs.27371	0	1	mRNA; cDNA DKFZp566J123 (from clone DKFZp566J1
54A11	709	1542	AK022811	Hs.27475	0	1	FLJ12749 fis, clone NT2RP2001149 /cds=UNK
111A5	42	686	NM_022485	Hs.27556	0	1	hypothetical protein FLJ22405 (FLJ22405), mR
123D4	879	1005	NM_016059	Hs.27693	3.00E-49	1	peptidylprolyl isomerase (cyclophilin)-like
518E11	1245	2235	AF332469	Hs.27721	0	5	putative protein WHSC1L1 (WHSC1L1) mRNA, comp
103B11	631	1343	NM_014805	Hs.28020	0	1	KIAA0766 gene product (KIAA0766), mRNA /cds=(
479H3	4	100	AB007928	Hs.28169	7.00E-37	1	mRNA for KIAA0459 protein, partial cds /cds=(0
526B3	1901	1995	NM_007218	Hs.28285	4.00E-47	1	patched related protein translocated in renal
480E4	4088	4596	AB046766	Hs.28338	0	1	mRNA for KIAA1546 protein, partial cds /cds=(0
164D10	651	970	NM_002970	Hs.28491	1.00E-163	2	spermidine/spermine N1-acetyltransferase (
69E10	729	1588	AB007888	Hs.28578	0	2	KIAA0428 mRNA, complete cds /cds=(1414,2526)
49B1	632	4266	NM_021038	Hs.28578	0	4	muscleblind (Drosophila)-like (MBNL), mRNA /

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

173A10	2105	2391	AL034548	Hs.28608	1.00E-161	2	DNA sequence from clone RP5-1103G7 on chromosome 20p1
156H8	467	585	AV691642	Hs.28739	8.00E-43	1	AV691642 5' end /clone=GKCDJG11 /clone_
588D3	444	909	NM_004800	Hs.28757	1.00E-123	1	transmembrane 9 superfamily member 2 (TM9SF2)
493B12	500	930	NM_003512	Hs.28777	0	1	H2A histone family, member L (H2AFL), mRNA /cd
115C5	63	661	BF341640	Hs.28788	0	1	602016073F1 cDNA, 5' end /clone=IMAGE:4151706
524C10	37	412	NM_007217	Hs.28866	1.00E-179	1	programmed cell death 10 (PDCD10), mRNA /cds=(
39A8	1380	1873	AK000196	Hs.29052	0	1	FLJ20189 fis, clone COLF0657 /cds=(122,84
477H7	690	1047	NM_005859	Hs.29117	1.00E-163	1	purine-rich element binding protein A (PURA),
134C8	2462	2789	NM_002894	Hs.29287	1.00E-173	1	retinoblastoma-binding protein 8 (RBBP8), mR
108A11	182	992	M31165	Hs.29352	0	9	tumor necrosis factor-inducible (TSG-6) mRNA fragme
99E8	179	992	NM_007115	Hs.29352	0	7	tumor necrosis factor, alpha-induced protein
169B3	2219	2683	AF039942	Hs.29417	0	1	HCF-binding transcription factor Zhangfei (Z
526A7	2219	2670	NM_021212	Hs.29417	0	1	HCF-binding transcription factor Zhangfei (Z
184H12	2380	4852	AB033042	Hs.29679	0	2	KIAA1216 protein, partial cds /cds=(0
125G9	1169	1814	AB037791	Hs.29716	0	1	mRNA for KIAA1370 protein, partial cds /cds=(4
68F3	1011	1892	AK027197	Hs.29797	0	5	FLJ23544 fis, clone LNG08336 /cds=(125,5
72H12	2103	2564	L27071	Hs.29877	0	2	tyrosine kinase (TXK) mRNA, complete cds /cds=(86,166
588D5	793	1321	NM_003328	Hs.29877	0	1	TXK tyrosine kinase (TXK), mRNA /cds=(86,1669)
127C3	1	1424	AK024961	Hs.29977	0	4	cDNA: FLJ21308 fis, clone COL02131 /cds=(287,1
128H7	351	977	NM_014188	Hs.30026	0	1	HSPC182 protein (HSPC182), mRNA /cds=(65,649)
521G4	502	1260	NM_004593	Hs.30035	0	4	splicing factor, arginine/serine-rich (trans
47A2	503	1265	U61267	Hs.30035	0	4	putative splice factor transformer2-beta mRN
37G9	1287	1763	M16967	Hs.30054	0	2	coagulation factor V mRNA, complete cds /cds=(90,6764
459E1	43	536	NM_015919	Hs.30303	0	1	Kruppel-associated box protein (LOC51595), m
465F6	256	573	NM_005710	Hs.30570	7.00E-75	1	polyglutamine binding protein 1 (PQBP1), mRNA
120H1	5305	5634	NM_012296	Hs.30687	1.00E-172	2	GRB2-associated binding protein 2 (GAB2), mRN
189G2	1	147	BG260954	Hs.30724	2.00E-68	1	602372562F1 cDNA, 5' end /clone=IMAGE:4480647
482E6	3086	3254	AK023743	Hs.30818	4.00E-91	1	cDNA FLJ13681 fis, clone PLACE2000014, weakly
179H5	20	1232	AK001972	Hs.30822	0	2	FLJ11110 fis, clone PLACE1005921, weakly
598B6	1	1169	NM_018326	Hs.30822	0	19	hypothetical protein FLJ11110 (FLJ11110), mR
126G10	1309	2463	AK000689	Hs.30882	0	18	cDNA FLJ20682 fis, clone KAIA3543, highly simi
126G7	5221	5904	NM_019081	Hs.30909	1.00E-163	2	KIAA0430 gene product (KIAA0430), mRNA /cds=(
483D1	1481	2098	NM_003098	Hs.31121	0	1	syntrophin, alpha 1(dystrophin-associated p
464C9	1188	1755	NM_003273	Hs.31130	0	1	transmembrane 7 superfamily member 2 (TM7SF2),
478A6	3024	3837	NM_012238	Hs.31176	1.00E-176	2	sir2-like 1 (SIRT1), mRNA /cds=(53,2296) /gb=
122E5	1060	1294	NM_002893	Hs.31314	1.00E-113	1	retinoblastoma-binding protein 7 (RBBP7), mR
117B1	2056	2489	AF153419	Hs.31323	0	1	IkappaBkinase complex-associated protein (I
462E10	337	569	AV752358	Hs.31409	1.00E-108	1	AV752358 cDNA, 5' end /clone=NPDBHG03 /clone_
126E7	1962	2748	AB014548	Hs.31921	0	2	mRNA for KIAA0648 protein, partial cds /cds=(0
186G11	729	954	BC000152	Hs.31989	1.00E-125	1	Similar to DKFZP586G1722 protein, clone MGC:
67H7	1705	2336	AJ400877	Hs.32017	0	2	ASCL3 gene, CEGP1 gene, C11orf14 gene, C11orf1
102B11	175	874	AK026455	Hs.32148	0	1	FLJ22802 fis, clone KAIA2682, highly sim
458D4	46	449	H14103	Hs.32149	1.00E-167	1	ym62a02.r1 cDNA, 5' end /clone=IMAGE:163466 /
99A2	3991	4532	AB007902	Hs.32168	0	1	KIAA0442 mRNA, partial cds /cds=(0,3519) /gb=
458G5	27	540	N30152	Hs.32250	0	1	yx81f03.s1 cDNA, 3' end /clone=IMAGE:268157 /
112D11	4399	5040	NM_005922	Hs.32353	0	1	mitogen-activated protein kinase kinase kina
48C8	3278	3988	AB002377	Hs.32556	0	2	mRNA for KIAA0379 protein, partial cds /cds=(0,
515F9	761	989	NM_003193	Hs.32675	1.00E-116	1	tubulin-specific chaperone e (TBCE), mRNA /c
158C12	342	809	NM_016063	Hs.32826	0	1	CGI-130 protein (LOC51020), /cds=(63,575
585E6	128	512	NM_005594	Hs.32916	0	3	nascent-polypeptide-associated complex alp
459B5	1271	1972	NM_017632	Hs.32922	0	1	hypothetical protein FLJ20036 (FLJ20036), mR

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

469G12	2711	2978	NM_001566	Hs.32944	1.00E-136	1	inositol polyphosphate-4-phosphatase, type
71B7	483	1787	NM_003037	Hs.32970	0	29	signaling lymphocytic activation molecule (S
74G1	1	1780	U33017	Hs.32970	0	33	signaling lymphocytic activation molecule (SLAM) mR
473B11	2993	3361	NM_006784	Hs.33085	1.00E-111	1	WD repeat domain 3 (WDR3), mRNA /cds=(47,2878)
56B5	23	578	AB019571	Hs.33190	0	1	expressed only in placental villi, clone
469D12	187	394	AL359654	Hs.33756	1.00E-110	1	mRNA full length insert cDNA clone EUROIMAGE 19
98H8	371	618	AI114652	Hs.33757	3.00E-98	1	HA1247 cDNA /gb=AI114652 /gi=6359997 /ug=Hs.
594E7	2134	2320	NM_012123	Hs.33979	5.00E-93	1	CGI-02 protein (CGI-02), mRNA /cds=(268,2124)
110D1	1158	1349	NM_018579	Hs.34401	1.00E-105	1	hypothetical protein PRO1278 (PRO1278), mRNA
596A6	1950	2144	NM_022766	Hs.34516	1.00E-102	2	hypothetical protein FLJ23239 (FLJ23239), mR
37B10	237	563	AI123826	Hs.34549	1.00E-145	1	ow61c10.x1 cDNA, 3' end /clone=IMAGE:1651314
458H4	3656	4415	AB040929	Hs.35089	0	1	mRNA for KIAA1496 protein, partial cds /cds=(0
100D1	3563	3777	D25215	Hs.35804	1.00E-105	1	KIAA0032 gene, complete cds /cds=(166,3318)
519A12	402	623	AW960004	Hs.36475	3.00E-48	1	EST372075 cDNA /gb=AW960004 /gi=8149688 /ug=
498H2	11143	11490	NM_000081	Hs.36508	0	1	Chediak-Higashi syndrome 1 (CHS1), mRNA /cds=(
521D6	304	791	NM_002712	Hs.36587	0	2	protein phosphatase 1, regulatory subunit 7 (
460E1	1200	1542	AF319476	Hs.36752	0	2	GKAP42 (FKSG21) mRNA, complete cds /cds=(174,1
184G9	498	1191	AF082569	Hs.36794	0	2	D-type cyclin-interacting protein 1 (DIP1) mR
462D3	493	1517	NM_012142	Hs.36794	0	3	D-type cyclin-interacting protein 1 (DIP1), m
74E12	659	3054	D86956	Hs.36927	0	23	KIAA0201 gene, complete cds /cds=(347,2923)
58G5	1268	2888	NM_006644	Hs.36927	0	12	heat shock 105kD (HSP105B), mRNA /cds=(313,275
52C10	1479	2588	AK022546	Hs.37747	0	2	FLJ12484 fis, clone NT2RM1001102, weakly
479F9	2066	2322	AL136932	Hs.37892	1.00E-119	1	mRNA; cDNA DKFZp586H1322 (from clone DKFZp586H
483C2	2222	2723	NM_003173	Hs.37936	0	1	suppressor of variegation 3-9 (Drosophila) ho
593G6	673	1213	NM_004510	Hs.38125	0	1	interferon-induced protein 75, 52kD (IFI75),
101G12	118	436	N39230	Hs.38218	1.00E-173	1	yy50c03.s1 cDNA, 3' end /clone=IMAGE:276964 /
107E5	238	525	AW188135	Hs.38664	1.00E-158	1	xj92g04.x1 cDNA, 3' end /clone=IMAGE:2664726
596F2	9	504	BF892532	Hs.38664	0	9	IL0-MT0152-061100-501-e04 cDNA /gb=BF892532
469D7	47	474	NM_014343	Hs.38738	0	1	claudin 15 (CLDN15), mRNA /cds=(254,940) /gb=
166H8	1	81	BF103848	Hs.39457	9.00E-34	1	601647352F1 cDNA, 5' end /clone=IMAGE:3931452
465F3	157	296	NM_017859	Hs.39850	2.00E-47	1	hypothetical protein FLJ20517 (FLJ20517), mR
195C12	2684	2944	NM_000885	Hs.40034	1.00E-146	1	integrin, alpha 4 (antigen CD49D, alpha 4 subu
151F11	1393	1661	AL031427	Hs.40094	6.00E-81	1	DNA sequence from clone 167A19 on chromosome 1p32.1-33
134C12	4532	4802	NM_004973	Hs.40154	1.00E-114	1	jumonji (mouse) homolog (JMJ), mRNA /cds=(244,
115C9	5279	5614	AB033085	Hs.40193	1.00E-157	1	mRNA for KIAA1259 protein, partial cds /cds=(1
119A8	862	2087	NM_006152	Hs.40202	0	3	lymphoid-restricted membrane protein (LRMP),
104D4	924	1398	U10485	Hs.40202	0	2	lymphoid-restricted membrane protein (Jaw1) mRNA, c
155G3	226	530	AF047472	Hs.40323	1.00E-114	1	spleen mitotic checkpoint BUB3 (BUB3) mRNA, c
521C2	233	710	NM_004725	Hs.40323	0	1	BUB3 (budding uninhibited by benzimidazoles 3
107B8	187	545	AI927454	Hs.40328	0	1	wo90a02.x1 cDNA, 3' end /clone=IMAGE:2462570
458F10	1	436	BE782824	Hs.40334	0	1	601472323F1 cDNA, 5' end /clone=IMAGE:3875501
463G6	16	496	AI266255	Hs.40411	0	1	qx69f01.x1 cDNA, 3' end /clone=IMAGE:2006617
162F1	2711	2895	D87468	Hs.40888	4.00E-96	1	KIAA0278 gene, partial cds /cds=(0,1383) /gb
463E1	70	272	AL137067	Hs.40919	1.00E-109	1	DNA sequence from clone RP11-13B9 on chromosome 9q22.
458E7	107	774	AK024474	Hs.41045	0	1	mRNA for FLJ00067 protein, partial cds /cds=(1
185G12	1051	2315	AL050141	Hs.41569	1.00E-140	11	mRNA; cDNA DKFZp586O031 (from clone DKFZp586O0
593F5	2106	2490	NM_006190	Hs.41694	0	1	origin recognition complex, subunit 2 (yeast h

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

513H4	739	1249	NM_002190	Hs.41724	0	6	interleukin 17 (cytotoxic T-lymphocyte-assoc
155F4	739	1247	U32659	Hs.41724	0	1	IL-17 mRNA, complete cds /cds=(53,520) /gb=U32659 /g
108H12	892	1227	L40377	Hs.41726	1.00E-170	1	cytoplasmic antiproteinase 2 (CAP2) mRNA, com
477E7	249	404	BG033294	Hs.41989	6.00E-75	1	602298548F1 cDNA, 5' end /clone=IMAGE:4393186
143E2	5775	6018	AB033112	Hs.42179	1.00E-136	2	for KIAA1286 protein, partial cds /cds=(1
586B10	720	1225	NM_001952	Hs.42287	0	1	E2F transcription factor 6 (E2F6), mRNA /cds=(
583A10	346	883	NM_012097	Hs.42500	0	1	ADP-ribosylation factor-like 5 (ARL5), mRNA
459A7	152	251	BC003525	Hs.42712	2.00E-50	1	Similar to Max, clone MGC:10775, mRNA, comple
37B7	43	2687	AF006082	Hs.42915	1.00E-130	2	actin-related protein Arp2 (ARP2) mRNA, compl
120E3	512	2426	NM_005722	Hs.42915	0	3	ARP2 (actin-related protein 2, yeast) homolog
99D1	3298	3761	NM_014939	Hs.42959	0	1	KIAA1012 protein (KIAA1012), mRNA /cds=(57,43
473B2	3025	3425	AK023647	Hs.43047	1.00E-164	1	cDNA FLJ13585 fis, clone PLACE1009150 /cds=UNK
460E6	2988	3184	AB033093	Hs.43141	1.00E-105	1	mRNA for KIAA1267 protein, partial cds /cds=(9
471F7	232	575	AW993524	Hs.43148	0	1	RC3-BN0034-120200-011-h06 cDNA /gb=AW993524
460B10	402	706	BE781009	Hs.43273	1.00E-78	1	601469768F1 cDNA, 5' end /clone=IMAGE:3872704
36F6	2815	3403	AK024439	Hs.43616	0	1	for FLJ00029 protein, partial cds /cds=(0
471G3	43	454	NM_006021	Hs.43628	1.00E-165	1	deleted in lymphocytic leukemia, 2 (DLEU2), mR
184H3	1819	2128	D14043	Hs.43910	1.00E-168	2	MGC-24, complete cds /cds=(79,648) /gb=D1404
195F4	511	2370	NM_006016	Hs.43910	0	7	CD164 antigen, sialomucin (CD164), mRNA /cds=
188H9	1573	2277	NM_006346	Hs.43913	0	3	PIBF1 gene product (PIBF1), mRNA /cds=(0,2276)
177H6	1575	2272	Y09631	Hs.43913	0	2	PIBF1 protein, complete /cds=(0,2276) /
481E6	2529	2873	AB032952	Hs.44087	1.00E-159	1	mRNA for KIAA1126 protein, partial cds /cds=(0
112F5	1105	1701	AF197569	Hs.44143	0	1	BAF180 (BAF180) mRNA, complete cds /cds=(96,48
146F5	2620	3147	AL117452	Hs.44155	0	1	DKFZp586G1517 (from clone DKFZp586G
514C5	166	431	NM_018838	Hs.44163	1.00E-149	3	13kDa differentiation-associated protein (L
71D9	1117	1800	AF263613	Hs.44198	0	2	membrane-associated calcium-independent ph
68E1	289	527	AA576946	Hs.44242	4.00E-83	1	nm82b03.s1 cDNA, 3' end /clone=IMAGE:1074701
53H12	1925	2112	X75042	Hs.44313	4.00E-84	1	rel proto-oncogene mRNA /cds=(177,2036) /gb=X75
595D4	21	402	NM_017867	Hs.44344	0	1	hypothetical protein FLJ20534 (FLJ20534), mR
165B10	250	658	BC000758	Hs.44468	0	1	clone MGC:2698, mRNA, complete cds /cds=(168,
592E9	37	2422	NM_002687	Hs.44499	0	5	pinin, desmosome associated protein (PNN), mR
69F10	14	1152	Y09703	Hs.44499	0	3	MEMA protein /cds=(406,2166) /gb=Y09703
458H6	1	352	NM_015697	Hs.44563	0	1	hypothetical protein (CL640), mRNA /cds=(0,39
182C11	690	1324	AB046861	Hs.44566	0	4	mRNA for KIAA1641 protein, partial cds /cds=(6
115G3	318	731	BG288837	Hs.44577	0	1	602388170F1 cDNA, 5' end /clone=IMAGE:4517129
70B11	1879	4363	U58334	Hs.44585	0	3	Bcl2, p53 binding protein Bbp/53BP2 (BBP/53BP2) mRNA
165F10	265	496	AV726117	Hs.44656	6.00E-66	1	AV726117 cDNA, 5' end /clone=HTCAXB05 /clone_
36F1	444	1176	AK001332	Hs.44672	0	1	FLJ10470 fis, clone NT2RP2000032, weakly
596H1	1073	2711	AF288571	Hs.44865	0	14	lymphoid enhancer factor-1 (LEF1) mRNA, compl
41C4	2876	3407	X60708	Hs.44926	0	1	pcHDP7 mRNA for liver dipeptidyl peptidase IV /cds=(75
588A7	7564	7849	AL031667	Hs.45207	1.00E-158	1	DNA sequence from clone RP4-620E11 on chromosome 20q1
183G6	3967	4942	AB020630	Hs.45719	0	5	mRNA for KIAA0823 protein, partial cds /cds=(0
465C9	700	1325	BC002796	Hs.46446	0	1	lymphoblastic leukemia derived sequence 1,
464B1	1519	1997	NM_006019	Hs.46465	0	1	T-cell, immune regulator 1 (TCIRG1), mRNA /cds
466F10	455	518	AW974756	Hs.46476	6.00E-26	1	EST386846 cDNA /gb=AW974756 /gi=8165944 /ug=
110E7	620	1153	AF223469	Hs.46847	0	1	AD022 protein (AD022) mRNA, complete cds /cds=
112D5	618	1197	NM_016614	Hs.46847	0	4	TRAF and TNF receptor-associated protein (AD0

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

172G6	4157	4527	NM_003954	Hs.47007	0	1	mitogen-activated protein kinase kinase kina
177C8	4217	4469	Y10256	Hs.47007	1.00E-96	1	serine/threonine protein kinase, NIK /c
458H9	18	457	AW291458	Hs.47325	0	1	UI-H-BI2-agh-c-02-0-UI.s1 cDNA, 3' end /clon
62B6	562	697	BE872760	Hs.47334	7.00E-54	1	601450902F1 cDNA, 5' end /clone=IMAGE:3854544
178F12	169	2413	AF307339	Hs.47783	0	2	B aggressive lymphoma short isoform (BAL) mRNA
460G4	598	1081	NM_005985	Hs.48029	0	1	snail 1 (drosophila homolog), zinc finger prot
70D12	1	2038	AK027070	Hs.48320	0	13	FLJ23417 fis, clone HEP20868 /cds=(59,12
41G5	6587	7128	NM_014345	Hs.48433	0	1	endocrine regulator (HRIHFB2436), mRNA /cds=
516H2	1	212	NM_017948	Hs.48712	2.00E-90	2	hypothetical protein FLJ20736 (FLJ20736), mR
517G9	665	1649	NM_004462	Hs.48876	0	2	farnesyl-diphosphate farnesyltransferase 1
146A2	88	440	X76770	Hs.49007	0	1	PAP /cds=UNKNOWN /gb=X76770 /gi=556782 /ug
174H4	2612	3200	AF189011	Hs.49163	0	1	ribonuclease III (RN3) mRNA, complete cds /cds
121G3	463	829	NM_017917	Hs.49376	0	1	hypothetical protein FLJ20644 (FLJ20644), mR
170B9	2260	2948	AK023825	Hs.49391	0	1	FLJ13763 fis, clone PLACE4000089 /cds=(56
65E2	629	1798	AF062075	Hs.49587	0	4	leupaxin mRNA, complete cds /cds=(93,1253) /g
518B2	26	1798	NM_004811	Hs.49587	0	12	leupaxin (LPXN), mRNA /cds=(93,1253) /gb=NM_0
472E8	1182	1516	AL390132	Hs.49822	0	1	mRNA; cDNA DKFZp547E107 (from clone DKFZp547E1
41B12	57	576	AB000887	Hs.50002	0	1	for EBI1-ligand chemokine, complete cds
41D1	1	310	U86358	Hs.50404	1.00E-135	1	chemokine (TECK) mRNA, complete cds /cds=(0,452) /gb
107C9	2861	3541	M64174	Hs.50651	0	3	protein-tyrosine kinase (JAK1) mRNA, complete cds /c
599H12	202	3541	NM_002227	Hs.50651	0	11	Janus kinase 1 (a protein tyrosine kinase) (JAK
105E3	621	1101	AF047442	Hs.50785	0	1	vesicle trafficking protein sec22b mRNA, comp
129B5	2489	2919	X16354	Hs.50964	0	2	transmembrane carcinoembryonic antigen BGPa
587H2	748	1673	NM_000521	Hs.51043	0	2	hexosaminidase B (beta polypeptide) (HEXB), m
458H12	4043	4561	NM_000887	Hs.51077	0	1	integrin, alpha X (antigen CD11C (p150), alpha
129C9	4055	4567	Y00093	Hs.51077	0	1	leukocyte adhesion glycoprotein p150,95
125D8	2502	3966	AF016266	Hs.51233	0	3	TRAIL receptor 2 mRNA, complete cds /cds=(117,1
179E1	17	343	M22538	Hs.51299	1.00E-179	1	nuclear-encoded mitochondrial NADH-ubiquinone redu
165D7	35	754	NM_021074	Hs.51299	0	4	NADH dehydrogenase (ubiquinone) flavoprotein
107F10	2632	2993	Y11251	Hs.51957	0	2	novel member of serine-arginine domain p
195B12	1344	1590	NM_017903	Hs.52184	3.00E-96	1	hypothetical protein FLJ20618 (FLJ20618), mR
69D7	3046	3568	AB014569	Hs.52526	0	4	for KIAA0669 protein, complete cds /cds=(
55D1	2607	2847	NM_014779	Hs.52526	1.00E-130	1	KIAA0669 gene product (KIAA0669), mRNA /cds=(
480B8	1943	2062	AL080213	Hs.52792	8.00E-44	1	mRNA; cDNA DKFZp586I1823 (from clone DKFZp586I
72G7	1236	1348	NM_018607	Hs.52891	2.00E-55	1	hypothetical protein PRO1853 (PRO1853), mRNA
526D1	1	256	NM_004597	Hs.53125	1.00E-114	1	small nuclear ribonucleoprotein D2 polypeptid
458E8	1182	1701	NM_002621	Hs.53155	0	1	properdin P factor, complement (PFC), mRNA /cd
458G2	2171	2836	NM_001204	Hs.53250	0	1	bone morphogenetic protein receptor, type II
458F7	30	650	NM_002200	Hs.54434	0	1	interferon regulatory factor 5 (IRF5), mRNA /
459F12	2023	3325	NM_006060	Hs.54452	0	2	zinc finger protein, subfamily 1A, 1 (Ikaros) (
41A6	498	755	U46573	Hs.54460	1.00E-140	1	eotaxin precursor mRNA, complete cds /cds=(53,346) /
590A10	243	659	NM_004688	Hs.54483	0	2	N-myc (and STAT) interactor (NMI), mRNA /cds=(
461C11	872	1415	NM_014291	Hs.54609	0	1	glycine C-acetyltransferase (2-amino-3-keto
170H5	412	1630	AJ243721	Hs.54642	0	3	for dTDP-4-keto-6-deoxy-D-glucose 4-re
521F5	270	1491	NM_013283	Hs.54642	0	8	methionine adenosyltransferase II, beta (MAT
189H5	737	1049	X76302	Hs.54649	1.00E-131	2	H.sapiens RY-1 mRNA for putative nucleic acid binding protei
599D10	2614	3035	AB029015	Hs.54886	0	5	mRNA for KIAA1092 protein, partial cds /cds=(0
458D5	1026	1676	AK027243	Hs.54890	0	1	cDNA: FLJ23590 fis, clone LNG14491 /cds=(709,1
37A10	1633	2040	AK026024	Hs.55024	0	1	FLJ22371 fis, clone HRC06680 /cds=(77,12
121A8	799	1217	NM_018053	Hs.55024	1.00E-160	1	hypothetical protein FLJ10307 (FLJ10307), mR

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

460B1	11195	11326	AF231023	Hs.55173	1.00E-45	1	protocadherin Flamingo 1 (FMI1) mRNA, complete
57F1	1450	2070	NM_003447	Hs.55481	0	2	zinc finger protein 165 (ZNF165), mRNA /cds=(5
68D10	979	2070	U78722	Hs.55481	0	4	zinc finger protein 165 (Zpf165) mRNA, complete
584G7	268	1674	NM_003753	Hs.55682	0	4	eukaryotic translation initiation factor 3,
161C8	63	394	NM_017897	Hs.55781	1.00E-177	1	hypothetical protein FLJ20604 (FLJ20604), mR
588F6	1	387	NM_016497	Hs.55847	0	1	hypothetical protein (LOC51258), mRNA /cds=(
597E10	334	2073	NM_004446	Hs.55921	0	5	glutamyl-prolyl-tRNA synthetase (EPRS), mRN
138H10	3603	4112	X54326	Hs.55921	0	1	glutaminyl-tRNA synthetase /cds=(58,43
121D5	3959	4192	AB018348	Hs.55947	1.00E-130	1	mRNA for KIAA0805 protein, partial cds /cds=(0
473D12	1428	1866	AJ245539	Hs.55968	0	2	partial mRNA for GalNAc-T5 (GALNT5 gene) /cds=
71E3	843	1724	NM_005542	Hs.56205	0	30	insulin induced gene 1 (INSIG1), mRNA /cds=(414
73F4	843	2495	U96876	Hs.56205	0	32	insulin induced protein 1 (INSIG1) gene, compl
75C8	180	2439	AJ277832	Hs.56247	0	13	for inducible T-cell co-stimulator (ICOS
187A6	2073	2255	AF195530	Hs.56542	2.00E-99	1	soluble aminopeptidase P (XPNPEP1) mRNA, comp
584H5	1496	1889	NM_001494	Hs.56845	1.00E-151	1	GDP dissociation inhibitor 2 (GDI2), mRNA /cds
460C5	2395	2860	AK022936	Hs.56847	0	1	cDNA FLJ12874 fis, clone NT2RP2003769 /cds=UNK
460B5	164	741	BC003581	Hs.56851	0	1	Similar to RIKEN cDNA 2900073H19 gene, clone
54G4	1359	1761	AK027232	Hs.57209	0	2	FLJ23579 fis, clone LNG13017 /cds=UNKNOW
192D8	1576	2872	AL136703	Hs.57209	0	3	mRNA; cDNA DKFZp566J091 (from clone DKFZp566J0
66F9	618	1056	U41654	Hs.57304	0	1	adenovirus protein E3-14.7k interacting protein 1 (
183A1	2093	2334	NM_003751	Hs.57783	1.00E-132	1	eukaryotic translation initiation factor 3,
117B3	6933	7225	NM_022898	Hs.57987	1.00E-154	3	B-cell lymphoma/leukaemia 11B (BCL11B), mRNA
74C11	273	359	BE739287	Hs.58066	7.00E-21	1	601556492F1 cDNA, 5' end /clone=IMAGE:3826247
174H2	5591	5977	AJ131693	Hs.58103	0	1	mRNA for AKAP450 protein /cds=(222,11948) /gb
599H8	26	993	NM_003756	Hs.58189	0	3	eukaryotic translation initiation factor 3,
168F12	295	593	U54559	Hs.58189	1.00E-166	1	translation initiation factor eIF3 p40 subuni
68B11	1	297	BE867841	Hs.58297	1.00E-146	1	601443614F1 cDNA, 5' end /clone=IMAGE:3847827
104A6	376	2578	AF001862	Hs.58435	0	3	FYN binding protein mRNA, complete cds /cds=(67
192E3	230	648	NM_001465	Hs.58435	0	4	FYN-binding protein (FYB-120/130) (FYB), mRN
73B4	1287	1763	AK022834	Hs.58488	0	1	FLJ12772 fis, clone NT2RP2001634, highly
100G3	1568	1786	NM_004850	Hs.58617	1.00E-108	1	Rho-associated, coiled-coil containing prot
116G9	1997	2464	NM_013352	Hs.58636	0	1	squamous cell carcinoma antigen recognized by
178C6	5	710	AV760147	Hs.58643	1.00E-111	5	AV760147 cDNA, 5' end /clone=MDSEPB12 /clone_
519B1	2203	2320	NM_014207	Hs.58685	1.00E-56	1	CD5 antigen (p56-62) (CD5), mRNA /cds=(72,1559
40B6	1655	2283	X04391	Hs.58685	0	1	lymphocyte glycoprotein T1/Leu-1 /cds=(72,1
466B9	262	534	AI684437	Hs.58774	1.00E-107	1	wa82a04.x1 cDNA, 3' end /clone=IMAGE:2302638
480H7	86	234	NM_006568	Hs.59106	1.00E-54	1	cell growth regulatory with ring finger domain
44A7	2229	2703	X17094	Hs.59242	0	1	fur mRNA for furin /cds=(216,2600) /gb=X17094 /gi=314
106D12	21	380	M96982	Hs.59271	0	2	U2 snRNP auxiliary factor small subunit, compl
39C5	1821	2653	AB011098	Hs.59403	0	1	for KIAA0526 protein, complete cds /cds=(
185H7	1826	2352	NM_004863	Hs.59403	0	1	serine palmitoyltransferase, long chain base
459C5	126	443	AA889552	Hs.59459	1.00E-158	1	ak20d12.s1 cDNA, 3' end /clone=IMAGE:1406519
108B8	2760	3079	AJ132592	Hs.59757	1.00E-138	1	for zinc finger protein, 3115 /cds=(107,27
194F7	2074	2461	NM_018227	Hs.59838	0	1	hypothetical protein FLJ10808 (FLJ10808), mR
465D4	2	132	AI440512	Hs.59844	7.00E-67	1	tc83f09.x1 cDNA, 3' end /clone=IMAGE:2072777
161H10	1	381	AA004799	Hs.60088	1.00E-169	1	zh96b05.s1 cDNA, 3' end /clone=IMAGE:429105 /
465B6	228	383	NM_018986	Hs.61053	1.00E-66	1	hypothetical protein (FLJ20356), mRNA /cds=(
102G9	359	725	D11094	Hs.61153	0	1	MSS1, complete cds /cds=(66,1367) /gb=D11094
193C6	359	725	NM_002803	Hs.61153	1.00E-174	2	proteasome (prosome, macropain) 26S subunit,
99E7	1768	2339	AL023653	Hs.61469	0	10	DNA sequence from clone 753P9 on chromosome Xq25-26.1.
462B9	5	411	BE779284	Hs.61472	1.00E-152	1	601464557F1 cDNA, 5' end /clone=IMAGE:3867566
594F11	220	569	NM_003905	Hs.61828	1.00E-159	2	amyloid beta precursor protein-binding prote

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

102E7	1216	1921	AF046001	Hs.62112	0	3	zinc finger transcription factor (ZNF207) mRN
192B4	754	934	NM_003457	Hs.62112	2.00E-98	2	zinc finger protein 207 (ZNF207), mRNA /cds=(2
41G9	1664	2096	J02931	Hs.62192	0	1	placental tissue factor (two forms) mRNA, complete cd
482E12	1857	2149	NM_001993	Hs.62192	5.00E-87	1	coagulation factor III (thromboplastin, tiss
459C10	1548	1845	AB011114	Hs.62209	1.00E-166	1	mRNA for KIAA0542 protein, partial cds /cds=(39
114D6	2251	2712	NM_002053	Hs.62661	0	1	guanylate binding protein 1, interferon-induc
590C9	83	760	NM_002032	Hs.62954	0	43	ferritin, heavy polypeptide 1 (FTH1), mRNA /c
458C5	1798	2407	AB033118	Hs.63128	0	1	mRNA for KIAA1292 protein, partial cds /cds=(0
109E5	4661	5114	AB002369	Hs.63302	0	1	KIAA0371 gene, complete cds /cds=(247,3843)
589G9	250	5650	NM_021090	Hs.63302	0	6	myotubularin related protein 3 (MTMR3), mRNA
182E4	1751	2144	NM_002831	Hs.63489	0	1	protein tyrosine phosphatase, non-receptor t
589C8	1787	2222	AK023529	Hs.63525	0	2	cDNA FLJ13467 fis, clone PLACE1003519, highly
458D7	1595	1912	NM_022727	Hs.63609	1.00E-180	1	HpaII tiny fragments locus 9C (HTF9C), mRNA /c
193A2	144	2588	NM_003264	Hs.63668	0	5	toll-like receptor 2 (TLR2), mRNA /cds=(129,24
117C3	1504	2366	AF131762	Hs.64001	0	3	clone 25218 mRNA sequence /cds=UNKNOWN /gb=AF
109F1	568	2157	AL031602	Hs.64239	0	3	DNA sequence from clone RP5-1174N9 on chromosome 1p34
40D5	698	1192	U32324	Hs.64310	0	1	interleukin-11 receptor alpha chain mRNA, complete c
522F4	12	504	NM_006356	Hs.64593	0	1	ATP synthase, H+ transporting, mitochondrial
462E9	215	891	NM_015423	Hs.64595	0	1	aminoadipate-semialdehyde dehydrogenase-ph
164G10	37	889	NM_006851	Hs.64639	0	2	glioma pathogenesis-related protein (RTVP1),
155G10	1	601	U16307	Hs.64639	0	1	glioma pathogenesis-related protein (GliPR) mRNA, c
110D11	341	712	S60099	Hs.64797	0	1	APPH=amyloid precursor protein homolog [human, placenta,
513E8	3411	3986	AF148537	Hs.65450	0	7	reticulon 4a mRNA, complete cds /cds=(141,3719
460F4	1415	1749	NM_018174	Hs.66048	1.00E-163	1	hypothetical protein FLJ10669 (FLJ10669), mR
478H8	486	1037	NM_001775	Hs.66052	0	1	CD38 antigen (p45) (CD38), mRNA /cds=(69,971)
461A6	2977	3516	AB051540	Hs.66053	0	1	mRNA for KIAA1753 protein, partial cds /cds=(0
191E7	1	494	AL157438	Hs.66151	0	6	mRNA; cDNA DKFZp434A115 (from clone DKFZp434A1
464B6	76	623	NM_002528	Hs.66196	0	1	nth (E.coli endonuclease III)-like 1 (NTHL1),
473C6	149	517	BE673759	Hs.66357	0	1	7d69d02.x1 cDNA, 3' end /clone=IMAGE:3278211
171G11	1001	1385	Z98884	Hs.66708	0	1	DNA sequence from clone RP3-467L1 on chromosome 1p36.
169H3	15	1800	X82200	Hs.68054	0	4	Staf50 /cds=(122,1450) /gb=X82200 /gi=8992
167G9	747	1104	NM_005932	Hs.68583	1.00E-101	1	mitochondrial intermediate peptidase (MIPEP)
170H3	747	1104	U80034	Hs.68583	6.00E-99	1	mitochondrial intermediate peptidase precurs
69F9	321	1348	U78027	Hs.69089	0	5	Bruton's tyrosine kinase (BTK), alpha-D-galac
586D6	16	676	NM_006360	Hs.69469	1.00E-173	2	dendritic cell protein (GA17), mRNA /cds=(51,1
591E3	74	189	NM_002385	Hs.69547	2.00E-59	1	myelin basic protein (MBP), mRNA /cds=(10,570)
597H2	482	2702	NM_007158	Hs.69855	0	8	NRAS-related gene (D1S155E), mRNA /cds=(420,2
515C5	3257	3421	NM_003169	Hs.70186	8.00E-45	1	suppressor of Ty (S.cerevisiae) 5 homolog (SUP
461B9	44	425	H06786	Hs.70258	0	1	yl83g05.r1 cDNA, 5' end /clone=IMAGE:44737 /c
525H4	2834	2978	NM_014933	Hs.70266	4.00E-77	1	yeast Sec31p homolog (KIAA0905), mRNA /cds=(53
521C3	1	1165	NM_016628	Hs.70333	1.00E-176	2	hypothetical protein (LOC51322), mRNA /cds=(
460E5	414	994	AF138903	Hs.70337	0	1	immunoglobulin superfamily protein beta-like
190C7	1406	1788	D50926	Hs.70359	0	1	mRNA for KIAA0136 gene, partial cds /cds=(0,2854) /gb
497F10	653	1096	NM_014210	Hs.70499	0	3	ecotropic viral integration site 2A (EVI2A), m
37C11	820	1523	AB002368	Hs.70500	0	4	KIAA0370 gene, partial cds /cds=(0,2406) /gb
464B2	496	721	BG283002	Hs.71243	3.00E-99	1	602406192F1 cDNA, 5' end /clone=IMAGE:4518214
69G4	1292	2708	AL161991	Hs.71252	0	4	cDNA DKFZp761C169 (from clone DKFZp761C1
485E4	176	485	AA131524	Hs.71433	1.00E-151	1	zl31h02.s1 cDNA, 3' end /clone=IMAGE:503571 /
161G2	1338	1877	NM_003129	Hs.71465	0	1	squalene epoxidase (SQLE), mRNA /cds=(214,193

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

188D6	328	597	NM_016630	Hs.71475	1.00E-129	1	hypothetical protein (LOC51324), mRNA /cds=(
483B5	12	384	NM_021128	Hs.71618	0	1	polymerase (RNA) II (DNA directed) polypeptide
161F6	675	1114	U79277	Hs.71848	0	1	clone 23548 mRNA sequence /cds=UNKNOWN /gb=U79277 /g
473F8	377	729	BE889075	Hs.71941	1.00E-146	1	601513514F1 cDNA, 5' end /clone=IMAGE:3915003
102A6	1129	1560	AK023183	Hs.72782	0	1	FLJ13121 fis, clone NT2RP3002687 /cds=(39
41E2	56	539	M57506	Hs.72918	0	1	secreted protein (I-309) gene, complete cds /cds=(72,
476E12	1790	2311	S76638	Hs.73090	0	2	p50-NF-kappa B homolog [human, peripheral blood T cells, mR
41G7	3116	3469	U64198	Hs.73165	1.00E-173	1	Il-12 receptor beta2 mRNA, complete cds /cds=(640,322
51C9	1721	2339	NM_005263	Hs.73172	0	4	growth factor independent 1 (GFI1), mRNA /cds=
67H6	1723	2342	U67369	Hs.73172	0	1	growth factor independence-1 (Gfi-1) mRNA, complete
179E7	211	610	M92444	Hs.73722	0	1	apurinic/apyrimidinic endonuclease (HAP1) g
585G3	174	589	NM_001641	Hs.73722	0	8	APEX nuclease (multifunctional DNA repair enz
138A11	1360	1717	M72709	Hs.73737	1.00E-151	1	alternative splicing factor mRNA, complete cds /cds=
49C8	1628	2276	AK001313	Hs.73742	0	4	cDNA FLJ10451 fis, clone NT2RP1000959, highly
41D7	2760	3563	J03565	Hs.73792	0	1	Epstein-Barr virus complement receptor type II(cr2)
121F8	2470	2815	AL136131	Hs.73793	1.00E-123	1	DNA sequence from clone RP1-261G23 on chromosome 6p12
482C7	2864	3199	NM_003005	Hs.73800	1.00E-165	3	selectin P (granule membrane protein 140kD, an
153E12	160	778	D90144	Hs.73817	0	22	gene for LD78 alpha precursor, complete cds /c
489E12	161	776	NM_002983	Hs.73817	0	6	small inducible cytokine A3 (homologous to mo
177D7	112	388	BF673951	Hs.73818	1.00E-143	1	602137331F1 cDNA, 5' end /clone=IMAGE:4274094
587E10	5	387	NM_006004	Hs.73818	1.00E-155	6	ubiquinol-cytochrome c reductase hinge prote
142H11	119	436	AL110183	Hs.73851	1.00E-148	1	cDNA DKFZp566A221 (from clone DKFZp566A2
190G11	1	375	NM_001685	Hs.73851	0	6	ATP synthase, H+ transporting, mitochondrial
119D10	675	1700	BC001267	Hs.73957	0	4	RAB5A, member RAS oncogene family, clone MGC:
135H12	1244	1772	NM_003016	Hs.73965	0	2	splicing factor, arginine/serine-rich 2 (SFR
160E6	1811	2196	X75755	Hs.73965	0	5	PR264 gene /cds=(98,763) /gb=X75755 /gi=455418
175F9	791	1446	L29218	Hs.73986	0	2	clk2 mRNA, complete cds /cds=(129,1628) /gb=L2
516D9	782	1144	NM_003992	Hs.73987	0	1	CDC-like kinase 3 (CLK3), transcript variant p
469F3	1778	1956	NM_002286	Hs.74011	4.00E-78	1	lymphocyte-activation gene 3 (LAG3), mRNA /cd
481D6	1323	1805	Z22970	Hs.74076	1.00E-173	1	H.sapiens mRNA for M130 antigen cytoplasmic variant 2 /cds=(
193H9	813	1569	NM_007360	Hs.74085	1.00E-127	3	DNA segment on chromosome 12 (unique) 2489 expr
39D9	810	994	X54870	Hs.74085	1.00E-100	1	NKG2-D gene /cds=(338,988) /gb=X54870 /gi=3
71F3	3014	3858	NM_004430	Hs.74088	1.00E-114	4	early growth response 3 (EGR3), mRNA /cds=(357,
74B12	3651	4214	S40832	Hs.74088	1.00E-114	7	EGR3=EGR3 protein mRNA,
105E11	2	142	AL050391	Hs.74122	6.00E-72	2	cDNA DKFZp586A181 (from clone DKFZp586A1
174A12	141	1072	NM_001225	Hs.74122	0	9	caspase 4, apoptosis-related cysteine protea
599E9	351	1864	AF279903	Hs.74267	0	6	60S ribosomal protein L15 (EC45) mRNA, complet
74F7	126	1867	AF283772	Hs.74267	0	8	clone TCBAP0781 mRNA sequence /cds=(40,654) /
156G12	554	831	AF034607	Hs.74276	1.00E-156	1	chloride channel ABP mRNA, complete cds /cds=(
118F4	1	148	BG112085	Hs.74313	7.00E-65	2	602283260F1 cDNA, 5' end /clone=IMAGE:4370727
70G10	1	2177	M16660	Hs.74335	0	26	90-kDa heat-shock protein gene, cDNA, complete cds /c
64D1	330	2219	NM_007355	Hs.74335	0	26	heat shock 90kD protein 1, beta (HSPCB), mRNA /
121E12	700	1033	NM_006826	Hs.74405	0	1	tyrosine 3-monooxygenase/tryptophan 5-monoo
177D3	480	1645	X57347	Hs.74405	0	2	HS1 protein /cds=(100,837) /gb=X57347 /

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

155A5	680	1176	U86602	Hs.74407	0	1	nucleolar protein p40 mRNA, complete cds /cds=(142,10
181G10	1802	2302	NM_012381	Hs.74420	0	2	origin recognition complex, subunit 3 (yeast h
66D8	927	1490	X86691	Hs.74441	0	1	218kD Mi-2 protein /cds=(89,5827) /gb=X
189D10	383	1102	NM_001749	Hs.74451	0	7	calpain 4, small subunit (30K) (CAPN4), mRNA /
171A3	721	1092	X04106	Hs.74451	1.00E-174	1	calcium dependent protease (small subunit) /
173F3	1069	1468	NM_004559	Hs.74497	0	1	nuclease sensitive element binding protein 1
176B7	1592	1990	NM_001178	Hs.74515	0	1	aryl hydrocarbon receptor nuclear translocato
481A11	2012	2210	NM_000947	Hs.74519	2.00E-61	1	primase, polypeptide 2A (58kD) (PRIM2A), mRNA
116G8	689	1417	NM_002537	Hs.74563	0	4	ornithine decarboxylase antizyme 2 (OAZ2), mR
526F6	185	1088	NM_003145	Hs.74564	0	3	signal sequence receptor, beta (translocon-as
104D3	713	1127	X79353	Hs.74576	0	1	XAP-4 mRNA for GDP-dissociation inhibitor /cds=(
518G1	2725	2993	NM_001357	Hs.74578	1.00E-134	1	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide
459H1	3093	3268	NM_014767	Hs.74583	3.00E-67	1	KIAA0275 gene product (KIAA0275), mRNA /cds=(
69C5	2304	2781	M97287	Hs.74592	0	3	MAR/SAR DNA binding protein (SATB1) mRNA
587F12	930	2777	NM_002971	Hs.74592	0	6	special AT-rich sequence binding protein 1 (b
124H10	1240	1812	NM_002808	Hs.74619	0	2	proteasome (prosome, macropain) 26S subunit,
57F10	700	2310	NM_000311	Hs.74621	0	60	prion protein (p27-30) (Creutzfeld-Jakob dis
74A10	870	2252	U29185	Hs.74621	0	34	prion protein (PrP) gene, complete cds /cds=(24
176H10	465	923	NM_000108	Hs.74635	0	1	dihydrolipoamide dehydrogenase (E3 component
98F4	870	2566	NM_003217	Hs.74637	0	7	testis enhanced gene transcript (TEGT), mRNA
179H8	1	1210	X75861	Hs.74637	0	3	TEGT gene /cds=(40,753) /gb=X75861 /gi=456258 /
125C4	417	1425	NM_014280	Hs.74711	0	2	splicing factor similar to dnaJ (SPF31), mRNA
74C5	21	177	BE549137	Hs.74861	4.00E-65	1	601076443F1 cDNA, 5' end /clone=IMAGE:3462154
497B12	124	384	NM_006713	Hs.74861	1.00E-123	2	activated RNA polymerase II transcription cof
191E10	497	859	NM_022451	Hs.74899	0	1	hypothetical protein FLJ12820 (FLJ12820), mR
114A3	1032	1446	AY007131	Hs.75061	0	1	clone CDABP0045 mRNA sequence
117G3	279	799	NM_004622	Hs.75066	0	1	translin (TSN), mRNA /cds=(81,767) /gb=NM_004
483G2	3293	3639	NM_006148	Hs.75080	1.00E-180	1	LIM and SH3 protein 1 (LASP1), /cds=(75,860) /g
181E11	8314	8804	NM_000038	Hs.75081	0	1	adenomatosis polyposis coli (APC), mRNA /cds=
597G6	374	2361	NM_003406	Hs.75103	0	6	tyrosine 3-monooxygenase/tryptophan 5-monoo
596F11	684	1088	NM_002097	Hs.75113	0	1	general transcription factor IIIA (GTF3A), mR
69C9	995	1564	AF113702	Hs.75117	0	4	clone FLC1353 PRO3063 mRNA, complete cds /cds=
46E7	128	1519	NM_004515	Hs.75117	1.00E-164	2	interleukin enhancer binding factor 2, 45kD (
481B10	66	515	NM_003201	Hs.75133	0	1	transcription factor 6-like 1 (mitochondrial
469C5	368	969	NM_006708	Hs.75207	0	1	glyoxalase I (GLO1), mRNA /cds=(87,641) /gb=N
71B4	939	2049	NM_002539	Hs.75212	0	24	ornithine decarboxylase 1 (ODC1) mRNA /cds=(33
75E10	173	1991	X16277	Hs.75212	0	51	ornithine decarboxylase ODC (EC 4.1.1.17) /c
166G9	2077	2632	L36870	Hs.75217	0	1	MAP kinase kinase 4 (MKK4) mRNA, complete cds /
167A12	2074	2619	NM_003010	Hs.75217	0	1	mitogen-activated protein kinase kinase 4 (M
105B12	3030	5207	D67029	Hs.75232	0	3	SEC14L mRNA, complete cds
125D1	4782	5209	NM_003003	Hs.75232	0	1	SEC14 (S. cerevisiae)-like 1 (SEC14L1), mRNA
184E4	2075	3174	D42040	Hs.75243	0	5	KIAA9001 gene, complete cds /cds=(1701,4106)
191E5	2071	3174	NM_005104	Hs.75243	0	2	bromodomain-containing 2 (BRD2), mRNA /cds=(1
186C12	4159	4866	NM_001068	Hs.75248	0	6	topoisomerase (DNA) II beta (180kD) (TOP2B), m
177C9	4473	4866	X68060	Hs.75248	0	1	topIIb mRNA for topoisomerase IIb /cds=(0,4865)
39D8	743	1980	D31885	Hs.75249	0	6	KIAA0069 gene, partial cds /cds=(0,680) /gb=
127G2	1363	1769	NM_016166	Hs.75251	0	1	DEAD/H (Asp-Glu-Ala-Asp/His) box binding pro
64E5	4	1214	NM_002922	Hs.75256	0	6	regulator of G-protein signalling 1 (RGS1), mR
69G5	276	914	S59049	Hs.75256	0	6	BL34=B cell activation gene [human, mRNA, 1398 nt]
101F6	315	758	AF054174	Hs.75258	0	1	histone macroH2A1.2 mRNA, complete cds /cds=(
596E10	320	1667	NM_004893	Hs.75258	0	5	H2A histone family, member Y (H2AFY), mRNA /cds
587G10	639	953	NM_001628	Hs.75313	1.00E-147	1	aldo-keto reductase family 1, member B1 (aldo

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

128F7	181	933	X06956	Hs.75318	0	4	HALPHA44 gene for alpha-tubulin, exons 1-3
74A1	321	3290	D21262	Hs.75337	0	10	KIAA0035 gene, partial cds /cds=(0,2125) /gb
50D8	2	667	BF303895	Hs.75344	0	4	601886515F2 cDNA, 5' end /clone=IMAGE:4120514
179F7	379	720	L07633	Hs.75348	1.00E-179	4	(clone 1950.2) interferon-gamma IEF SSP 5111 m
191F3	158	872	NM_006263	Hs.75348	0	18	proteasome (prosome, macropain) activator su
463G4	1849	2394	NM_001873	Hs.75360	0	1	carboxypeptidase E (CPE), mRNA /cds=(290,1720
117D6	224	671	AB023200	Hs.75361	0	1	mRNA for KIAA0983 protein, complete cds /cds=(
73E8	1	2339	D89077	Hs.75367	0	8	for Src-like adapter protein, complete cd
49H5	1	2388	NM_006748	Hs.75367	0	4	Src-like-adapter (SLA), mRNA /cds=(41,871) /
134A3	550	1126	NM_005917	Hs.75375	0	1	malate dehydrogenase 1, NAD (soluble) (MDH1),
462F2	73	361	NM_004172	Hs.75379	1.00E-158	1	solute carrier family 1 (glial high affinity gl
477G6	769	2043	NM_004300	Hs.75393	0	3	acid phosphatase 1, soluble (ACP1), transcript
62A10	1028	2528	X87949	Hs.75410	0	7	BiP protein /cds=(222,2183) /gb=X87949
125H4	510	807	NM_006010	Hs.75412	1.00E-130	2	Arginine-rich protein (ARP), mRNA /cds=(132,8
70H1	29	2349	AK026463	Hs.75415	0	30	FLJ22810 fis, clone KAIA2933, highly sim
60D3	160	1666	D31767	Hs.75416	0	6	KIAA0058 gene, complete cds /cds=(69,575) /g
98D5	103	1233	NM_014764	Hs.75416	0	10	DAZ associated protein 2 (DAZAP2), mRNA /cds=(
55H1	1183	1390	NM_016525	Hs.75425	2.00E-81	1	ubiquitin associated protein (UBAP), mRNA /cd
44B12	51	480	BF131654	Hs.75428	0	3	601820480F1 cDNA, 5' end /clone=IMAGE:4052586
64E11	1	177	NM_000454	Hs.75428	7.00E-94	1	superoxide dismutase 1, soluble (amyotrophic
65D3	387	969	L33842	Hs.75432	0	4	(clone FFE-7) type II inosine monophosphate de
58F9	379	672	NM_000884	Hs.75432	1.00E-149	1	IMP (inosine monophosphate) dehydrogenase 2
73B1	87	291	BE790474	Hs.75458	5.00E-71	2	601476059F1 cDNA, 5' end /clone=IMAGE:3878799
585G5	1	302	NM_000979	Hs.75458	1.00E-170	8	ribosomal protein L18 (RPL18), mRNA /cds=(15,5
173A1	1893	2653	NM_006763	Hs.75462	0	2	BTG family, member 2 (BTG2), mRNA /cds=(71,547)
166A10	601	1147	AB000115	Hs.75470	0	1	mRNA expressed in osteoblast, complete cds /cd
180D10	601	1045	NM_006820	Hs.75470	0	1	hypothetical protein, expressed in osteoblast
122D9	3322	5191	AB023173	Hs.75478	0	2	mRNA for KIAA0956 protein, partial cds /cds=(0
461E5	2484	2804	AL133074	Hs.75497	1.00E-144	1	mRNA; cDNA DKFZp434M1317 (from clone DKFZp434M
512D6	69	799	NM_004591	Hs.75498	0	12	small inducible cytokine subfamily A (Cys-Cys
146B12	54	783	U64197	Hs.75498	0	4	chemokine exodus-1 mRNA, complete cds /cds=(4
596H5	685	1952	NM_001157	Hs.75510	0	5	annexin A11 (ANXA11), mRNA /cds=(178,1695) /g
179D6	215	603	D23662	Hs.75512	1.00E-168	2	ubiquitin-like protein, complete cds
522G12	52	603	NM_006156	Hs.75512	0	2	neural precursor cell expressed, developmenta
46B6	1108	1418	NM_000270	Hs.75514	1.00E-166	1	nucleoside phosphorylase (NP), mRNA /cds=(109
73H11	83	1418	X00737	Hs.75514	1.00E-104	3	purine nucleoside phosphorylase (PNP; EC 2.
154F7	1279	2056	L05425	Hs.75528	0	3	nucleolar GTPase mRNA, complete cds /cds=(79,2
164C10	1268	1910	NM_013285	Hs.75528	0	2	nucleolar GTPase (HUMAUANTIG), mRNA /cds=(79,
106C8	76	322	Z25749	Hs.75538	1.00E-130	3	gene for ribosomal protein S7 /cds=(81,665) /gb=
98E5	474	1188	NM_003405	Hs.75544	0	1	tyrosine 3-monooxygenase/tryptophan 5-monoo
459G10	2160	2717	NM_000418	Hs.75545	0	1	interleukin 4 receptor (IL4R), mRNA /cds=(175,
44B2	71	692	U03851	Hs.75546	0	1	capping protein alpha mRNA, partial cds /cds=(16,870)
483F2	1207	1392	NM_004357	Hs.75564	1.00E-80	1	CD151 antigen (CD151), mRNA /cds=(84,845) /gb
596D6	1968	2392	NM_021975	Hs.75569	0	1	v-rel avian reticuloendotheliosis viral onco
466G10	679	896	NM_014763	Hs.75574	1.00E-120	2	mitochondrial ribosomal protein L19 (MRPL19),
524B3	6194	6477	NM_001759	Hs.75586	1.00E-147	1	cyclin D2 (CCND2), mRNA /cds=(269,1138) /gb=N
481B4	3423	3804	NM_000878	Hs.75596	1.00E-160	2	interleukin 2 receptor, beta (IL2RB), mRNA /cd
162B5	753	1694	M29064	Hs.75598	0	6	hnRNP B1 protein mRNA /cds=(149,1210) /gb=M29064 /gi
176F5	730	922	NM_002137	Hs.75598	1.00E-106	1	heterogeneous nuclear ribonucleoprotein A2/
106C2	1654	2589	D10522	Hs.75607	0	8	for 80K-L protein, complete cds /cds=(369,
98C5	1538	2589	NM_002356	Hs.75607	0	20	myristoylated alanine-rich protein kinase C

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

192E5	1007	1416	NM_006819	Hs.75612	0	1	stress-induced-phosphoprotein 1 (Hsp70/Hsp9
40E12	836	1765	M98399	Hs.75613	0	2	antigen CD36 (clone 21) mRNA, complete cds /cds=(254,1
107C6	1491	1595	AF113676	Hs.75621	3.00E-51	1	clone FLB2803 PRO0684 mRNA, complete cds /cds=
117E9	149	1033	NM_001779	Hs.75626	0	2	CD58 antigen, (lymphocyte function-associate
482H10	740	1367	NM_000591	Hs.75627	0	1	CD14 antigen (CD14), mRNA /cds=(119,1246) /gb
482D4	1342	1659	NM_006163	Hs.75643	3.00E-82	1	nuclear factor (erythroid-derived 2), 45kD (N
73F8	2864	3657	L49169	Hs.75678	0	20	G0S3 mRNA, complete cds /cds=(593,1609) /gb=L49169 /
58G3	3222	3657	NM_006732	Hs.75678	0	6	FBJ murine osteosarcoma viral oncogene homolo
53A7	30	836	J04130	Hs.75703	0	138	activation (Act-2) mRNA, complete cds /cds=(108,386)
500E11	41	688	NM_002984	Hs.75703	0	128	small inducible cytokine A4 (homologous to mo
170E9	415	2376	M16985	Hs.75709	0	6	cation-dependent mannose 6-phosphate-specific rece
591E8	1759	2401	NM_002355	Hs.75709	0	3	mannose-6-phosphate receptor (cation depende
191A11	20	1900	NM_002575	Hs.75716	0	13	serine (or cysteine) proteinase inhibitor, cl
184F5	18	1900	Y00630	Hs.75716	0	8	Arg-Serpin (plasminogen activator-inhibito
593G8	238	747	NM_005022	Hs.75721	1.00E-110	2	profilin 1 (PFN1), mRNA /cds=(127,549) /gb=NM
178G9	504	2101	NM_002951	Hs.75722	0	2	ribophorin II (RPN2), mRNA /cds=(288,2183) /g
138F12	2341	2488	Y00282	Hs.75722	4.00E-60	1	ribophorin II /cds=(288,2183) /gb=Y00282 /g
37F7	1328	1863	AK023290	Hs.75748	0	3	FLJ13228 fis, clone OVARC1000085, highly
119C7	3736	4103	NM_003137	Hs.75761	1.00E-172	1	SFRS protein kinase 1 (SRPK1), mRNA /cds=(108,2
52E8	574	1106	M36820	Hs.75765	0	2	cytokine (GRO-beta) mRNA, complete cds /cds=(74,397)
74C8	2055	3026	M10901	Hs.75772	0	4	glucocorticoid receptor alpha mRNA, complete cds /cd
196C5	2600	4591	NM_000176	Hs.75772	0	5	nuclear receptor subfamily 3, group C, member
68E7	2194	2597	D87953	Hs.75789	0	1	RTP, complete cds /cds=(122,1306) /gb=D87953
116E3	289	621	NM_016470	Hs.75798	0	1	hypothetical protein (HSPC207), mRNA /cds=(0
107C10	650	1165	AK025732	Hs.75811	0	1	FLJ22079 fis, clone HEP13180, highly sim
123C12	459	969	NM_004315	Hs.75811	0	1	N-acylsphingosine amidohydrolase (acid cera
99E11	1007	2346	NM_014761	Hs.75824	0	2	KIAA0174 gene product (KIAA0174), mRNA /cds=(
128C11	377	906	NM_006817	Hs.75841	0	2	endoplasmic reticulum lumenal protein (ERP28
175F5	455	843	X94910	Hs.75841	1.00E-173	1	ERp28 protein /cds=(11,796) /gb=X9491
182F12	4263	4842	D86550	Hs.75842	0	1	mRNA for serine/threonine protein kinase, complete c
175E3	3255	3787	AL110132	Hs.75875	0	1	mRNA; cDNA DKFZp564H192 (from clone DKFZp564H1
195G3	1435	2132	NM_003349	Hs.75875	0	2	ubiquitin-conjugating enzyme E2 variant 1 (U
184B12	17	282	BF698920	Hs.75879	1.00E-138	8	602126495F1 cDNA, 5' end /clone=IMAGE:4283350
67G6	1218	1605	AK000639	Hs.75884	1.00E-173	1	FLJ20632 fis, clone KAT03756, highly simi
516A11	721	1109	NM_015416	Hs.75884	0	2	DKFZP586A011 protein (DKFZP586A011), mRNA /c
44B1	1066	4914	NM_004371	Hs.75887	0	4	coatomer protein complex, subunit alpha (COPA
594D3	3971	4158	NM_003791	Hs.75890	1.00E-73	1	site-1 protease (subtilisin-like, sterol-reg
459H8	5291	5688	D87446	Hs.75912	1.00E-160	1	mRNA for KIAA0257 gene, partial cds /cds=(0,5418) /gb
113F6	2281	2807	NM_006842	Hs.75916	0	1	splicing factor 3b, subunit 2, 145kD (SF3B2), m
104F9	2334	2804	U41371	Hs.75916	0	1	spliceosome associated protein (SAP 145) mRNA, compl
100F12	656	825	AK024890	Hs.75932	6.00E-83	1	FLJ21237 fis, clone COL01114 /cds=UNKNOW
39E1	40	526	BF217687	Hs.75968	1.00E-124	2	601882510F1 cDNA, 5' end /clone=IMAGE:4094907
111G8	41	547	NM_021109	Hs.75968	1.00E-166	19	thymosin, beta 4, X chromosome (TMSB4X), mRNA
478A7	1335	1653	NM_006813	Hs.75969	1.00E-119	1	proline-rich protein with nuclear targeting s

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

70E9	652	1065	U03105	Hs.75969	0	1	B4-2 protein mRNA, complete cds /cds=(113,1096) /gb=U
596B9	508	1461	NM_003133	Hs.75975	0	2	signal recognition particle 9kD (SRP9), mRNA
513F12	1359	2169	NM_005151	Hs.75981	0	3	ubiquitin specific protease 14 (tRNA-guanine
74B3	1361	2166	U30888	Hs.75981	0	2	tRNA-guanine transglycosylase mRNA, complete cds /c
67B6	81	1457	X17025	Hs.76038	0	4	homolog of yeast IPP isomerase /cds=(50,736) /gb=X170
586F2	1471	2197	NM_004396	Hs.76053	0	13	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide
70B3	762	2211	X52104	Hs.76053	0	12	p68 protein /cds=(175,2019) /gb=X52104 /gi=3
73B2	32	494	BF214146	Hs.76064	0	1	601847762F1 cDNA, 5' end /clone=IMAGE:4078622
523E6	10	441	NM_000990	Hs.76064	0	2	ribosomal protein L27a (RPL27A), mRNA /cds=(1
38F7	6	372	Z23090	Hs.76067	0	2	28 kDa heat shock protein /cds=(491,1108)
59B6	916	1274	AF071596	Hs.76095	1.00E-174	1	apoptosis inhibitor (IEX-1L) gene, complete c
493B3	540	1206	NM_003897	Hs.76095	0	3	immediate early response 3 (IER3), mRNA /cds=(
483D7	1399	2063	NM_005626	Hs.76122	0	1	splicing factor, arginine/serine-rich 4 (SFR
591C12	13412	13873	NM_003922	Hs.76127	0	3	hect (homologous to the E6-AP (UBE3A) carboxyl
65H7	12209	12580	U50078	Hs.76127	0	1	guanine nucleotide exchange factor p532 mRNA, complet
160B6	79	535	X77584	Hs.76136	1.00E-140	1	ATL-derived factor/thiredoxin /cds=(80
596A9	1	124	NM_001009	Hs.76194	3.00E-62	1	ribosomal protein S5 (RPS5), mRNA /cds=(37,651
51H5	2834	3174	AK025353	Hs.76230	1.00E-180	1	cDNA: FLJ21700 fis, clone COL09849, highly sim
115C8	1589	2005	NM_001748	Hs.76288	0	1	calpain 2, (m/II) large subunit (CAPN2), mRNA
588C5	4	336	NM_004492	Hs.76362	0	2	general transcription factor IIA, 2 (12kD subu
111D9	732	1077	NM_004930	Hs.76368	1.00E-161	2	capping protein (actin filament) muscle Z-lin
192A11	1589	1995	NM_002462	Hs.76391	0	3	myxovirus (influenza) resistance 1, homolog o
39F5	8481	8730	Y00285	Hs.76473	1.00E-111	1	insuline-like growth factor II receptor /cds
98C4	487	3719	NM_002298	Hs.76506	0	38	lymphocyte cytosolic protein 1 (L-plastin) (L
124H12	611	1747	NM_004862	Hs.76507	0	5	LPS-induced TNF-alpha factor (PIG7), mRNA /cd
37A6	920	1524	U77396	Hs.76507	1.00E-162	2	LPS-Induced TNF-Alpha Factor (LITAF) mRNA, co
71E9	759	3362	D00099	Hs.76549	0	4	mRNA for Na,K-ATPase alpha-subunit, complete
73F5	951	1277	AK001361	Hs.76556	1.00E-168	1	FLJ10499 fis, clone NT2RP2000346, weakly
48H6	1097	1603	NM_014330	Hs.76556	0	2	growth arrest and DNA-damage-inducible 34 (G
160C8	74	181	BE730376	Hs.76572	2.00E-40	1	601563816F1 5' end /clone=IMAGE:3833690
589D11	86	455	NM_001697	Hs.76572	0	2	ATP synthase, H+ transporting, mitochondrial
38B1	227	886	NM_014059	Hs.76640	0	9	RGC32 protein (RGC32), mRNA /cds=(146,499) /g
174B12	3024	4628	D80005	Hs.76666	1.00E-136	4	mRNA for KIAA0183 gene, partial cds /cds=(0,3190) /gb
37A11	1788	3255	AF070673	Hs.76691	0	5	stannin mRNA, complete cds /cds=(175,441) /gb
58H11	1706	2088	AL136807	Hs.76698	0	2	mRNA; cDNA DKFZp434L1621 (from clone DKFZp434L
477F9	6930	7298	AB002299	Hs.76730	0	2	mRNA for KIAA0301 gene, partial cds /cds=(0,6144) /gb
40G7	293	819	NM_000118	Hs.76753	0	1	endoglin (Osler-Rendu-Weber syndrome 1) (EN
75C11	10	1113	J00194	Hs.76807	0	5	human hla-dr antigen alpha-chain mrna & ivs fragments /cds=
99F4	10	969	NM_019111	Hs.76807	0	6	major histocompatibility complex, class II,
61G12	1870	2511	AL133096	Hs.76853	0	1	cDNA DKFZp434N1728 (from clone DKFZp434N
599C2	41	346	NM_002790	Hs.76913	1.00E-124	1	proteasome (prosome, macropain) subunit, alp
155C2	508	870	X61970	Hs.76913	0	1	for macropain subunit zeta /cds=(21,746) /g
70C5	3398	3754	AF002020	Hs.76918	0	1	Niemann-Pick C disease protein (NPC1) mRNA, co
57A11	2173	2764	NM_000271	Hs.76918	0	1	Niemann-Pick disease, type C1 (NPC1), mRNA /cd
158C9	314	1233	NM_001679	Hs.76941	0	3	ATPase, Na+/K+ transporting, beta 3 polypeptid
520E1	4175	4502	NM_014757	Hs.76986	1.00E-158	1	mastermind (Drosophila), homolog of (MAML1),
587D8	22	869	NM_001006	Hs.77039	0	5	ribosomal protein S3A (RPS3A), mRNA /cds=(36,8
481F2	440	1488	NM_001731	Hs.77054	0	3	B-cell translocation gene 1, anti-proliferati
53G11	340	1490	X61123	Hs.77054	0	3	BTG1 mRNA /cds=(308,823) /gb=X61123 /gi=29508 /ug=Hs
521A6	147	1325	D55716	Hs.77152	0	2	mRNA for P1cdc47, complete cds /cds=(116,2275) /gb=D

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

37H9	2109	2530	X07109	Hs.77202	0	1	protein kinase C (PKC) type /cds=(136,2157) /
167H5	3915	4508	NM_006437	Hs.77225	0	1	ADP-ribosyltransferase (NAD+; poly (ADP-ribo
139G5	2183	2389	U61145	Hs.77256	1.00E-111	1	enhancer of zeste homolog 2 (EZH2) mRNA, complete cds
109H2	2502	2893	D38549	Hs.77257	0	1	KIAA0068 gene, partial cds /cds=(0,3816) /gb
184B7	619	1111	L25080	Hs.77273	0	1	GTP-binding protein (rhoA) mRNA, complete cds
587H1	614	1371	NM_001664	Hs.77273	0	9	ras homolog gene family, member A (ARHA), mRNA
99G10	1387	2219	NM_002658	Hs.77274	0	1	plasminogen activator, urokinase (PLAU), mRN
143C12	2403	2905	AL049332	Hs.77311	0	2	cDNA DKFZp564L176 (from clone DKFZp564L1
519B11	5248	5555	NM_000430	Hs.77318	1.00E-160	1	platelet-activating factor acetylhydrolase,
52F10	3249	3459	AF095901	Hs.77324	1.00E-114	2	eRF1 gene, complete cds /cds=(136,1449) /gb=A
494G1	3255	3453	NM_004730	Hs.77324	1.00E-109	2	eukaryotic translation termination factor 1
517E4	305	973	NM_014754	Hs.77329	0	2	phosphatidylserine synthase 1 (PTDSS1), mRNA
72F9	1934	4605	AF187320	Hs.77356	0	10	transferrin receptor (TFRC) gene, complete cd
46D6	241	4902	NM_003234	Hs.77356	0	2	transferrin receptor (p90, CD71) (TFRC), mRNA
113A12	1028	1290	NM_024033	Hs.77365	1.00E-145	1	hypothetical protein MGC5242 (MGC5242), mRNA
173A7	1142	1649	AK026164	Hs.77385	0	2	cDNA: FLJ22511 fis, clone HRC11837, highly sim
189E7	466	798	NM_002004	Hs.77393	0	1	farnesyl diphosphate synthase (farnesyl pyro
479B1	306	482	NM_000566	Hs.77424	8.00E-55	1	Fc fragment of IgG, high affinity Ia, receptor
41E12	351	898	X14356	Hs.77424	0	1	high affinity Fc receptor (FcRI) /cds=(36,116
122D3	562	855	NM_002664	Hs.77436	1.00E-145	1	pleckstrin (PLEK), mRNA /cds=(60,1112) /gb=N
59C11	1	2745	X07743	Hs.77436	0	5	pleckstrin (P47) /cds=(60,1112) /gb=X07743
590B1	5185	5274	NM_001379	Hs.77462	1.00E-44	1	DNA (cytosine-5-)-methyltransferase 1 (DNMT1
522D1	572	956	NM_001929	Hs.77494	0	1	deoxyguanosine kinase (DGUOK), mRNA /cds=(11,
109E12	723	2474	D87684	Hs.77495	1.00E-163	5	for KIAA0242 protein, partial cds /cds=(0,
148E2	61	271	BE737246	Hs.77496	1.00E-81	1	601305556F1 5' end /clone=IMAGE:3640165
586D4	1887	2362	NM_003363	Hs.77500	0	1	ubiquitin specific protease 4 (proto-oncogene
57E8	29	2808	BC001854	Hs.77502	0	30	methionine adenosyltransferase II, alpha, c
70H9	87	1283	X68836	Hs.77502	0	14	S-adenosylmethionine synthetase /cds=(
69B2	778	3033	M20867	Hs.77508	0	2	glutamate dehydrogenase (GDH) mRNA, complete cds /cd
513F9	2694	2929	NM_005271	Hs.77508	1.00E-105	1	glutamate dehydrogenase 1 (GLUD1), mRNA /cds=
75A3	190	701	X62744	Hs.77522	0	1	RING6 mRNA for HLA class II alpha product /cds=(45,830
105E10	72	597	BE673364	Hs.77542	0	3	7d34a03.x1 cDNA, 3' end /clone=IMAGE:3249100
124B2	85	683	BF508702	Hs.77542	0	8	UI-H-BI4-aop-g-05-0-UI.s1 cDNA, 3' end /clon
524C9	829	1233	AK021563	Hs.77558	0	3	cDNA FLJ11501 fis, clone HEMBA1002100 /cds=UNK
523B12	7580	8153	NM_004652	Hs.77578	0	2	ubiquitin specific protease 9, X chromosome (D
166F3	169	340	AL021546	Hs.77608	7.00E-63	1	DNA sequence from BAC 15E1 on chromosome 12. Contains
195A11	164	451	NM_003769	Hs.77608	1.00E-162	1	splicing factor, arginine/serine-rich 9 (SF
595E1	618	1461	AF056322	Hs.77617	0	7	SP100-HMG nuclear autoantigen (SP100) mRNA, c
115A6	2954	3541	AL137938	Hs.77646	0	2	mRNA; cDNA DKFZp761M0223 (from clone DKFZp761M
592H6	261	951	NM_014752	Hs.77665	0	3	KIAA0102 gene product (KIAA0102), mRNA /cds=(
461F3	4657	4980	NM_014749	Hs.77724	1.00E-174	1	KIAA0586 gene product (KIAA0586), mRNA /cds=(
98C8	27	1961	NM_002543	Hs.77729	0	4	oxidised low density lipoprotein (lectin-like
598A12	101	1396	NM_006759	Hs.77837	0	4	UDP-glucose pyrophosphorylase 2 (UGP2), mRNA
594H8	1	872	NM_006802	Hs.77897	1.00E-144	2	splicing factor 3a, subunit 3, 60kD (SF3A3), mR
171E4	1140	1394	X81789	Hs.77897	1.00E-110	1	for splicing factor SF3a60 /cds=(565,2070)
500F1	2185	2496	AK025736	Hs.77910	1.00E-160	1	cDNA: FLJ22083 fis, clone HEP14459, highly sim
525B10	1696	2060	NM_000122	Hs.77929	0	1	excision repair cross-complementing rodent r
53E1	877	1539	AK026595	Hs.77961	0	7	FLJ22942 fis, clone KAT08170, highly sim
521C6	631	1089	NM_005514	Hs.77961	1.00E-115	4	major histocompatibility complex, class I, B
588C3	300	653	NM_004792	Hs.77965	0	1	Clk-associating RS-cyclophilin (CYP), mRNA

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

523C6	277	582	NM_001912	Hs.78056	1.00E-143	1	cathepsin L (CTSL), mRNA /cds=(288,1289) /gb=
140D10	292	1549	X12451	Hs.78056	0	3	pro-cathepsin L (major excreted protein MEP)
463E5	129	552	NM_005969	Hs.78103	0	1	nucleosome assembly protein 1-like 4 (NAP1L4)
166H3	540	895	U77456	Hs.78103	0	1	nucleosome assembly protein 2 mRNA, complete cds /cd
40B10	2433	2543	M28526	Hs.78146	5.00E-29	1	platelet endothelial cell adhesion molecule (PECAM-1
114E5	1671	2029	NM_000442	Hs.78146	1.00E-162	1	platelet/endothelial cell adhesion molecule
513D11	28	1399	NM_000700	Hs.78225	0	5	annexin A1 (ANXA1), mRNA /cds=(74,1114) /gb=N
331B3	219	1370	X05908	Hs.78225	0	3	lipocortin /cds=(74,1114) /gb=X05908 /gi=34
56A12	1383	2379	X94232	Hs.78335	0	4	novel T-cell activation protein /cds=(14
465H1	386	904	NM_002812	Hs.78466	0	2	proteasome (prosome, macropain) 26S subunit,
108H7	2067	2486	L42572	Hs.78504	0	1	p87/89 gene, complete cds /cds=(92,2368) /gb=
187E9	729	1494	NM_006839	Hs.78504	0	2	inner membrane protein, mitochondrial (mitofi
102F2	672	2947	L14561	Hs.78546	0	2	plasma membrane calcium ATPase isoform 1 (ATP
591H12	42	1949	NM_004034	Hs.78637	0	3	annexin A7 (ANXA7), transcript variant 2, mRN
595H3	2775	3030	NM_003470	Hs.78683	3.00E-96	1	ubiquitin specific protease 7 (herpes virus-as
62F5	2775	3838	Z72499	Hs.78683	0	2	herpesvirus associated ubiquitin-speci
46G4	2632	3238	NM_003580	Hs.78687	0	1	neutral sphingomyelinase (N-SMase) activatio
513A11	342	1258	NM_002635	Hs.78713	0	10	solute carrier family 25 (mitochondrial carri
472A4	3018	3286	NM_024298	Hs.78768	1.00E-132	1	malignant cell expression-enhanced gene/tumo
177A3	377	1186	AL049589	Hs.78771	0	3	DNA sequence from clone 570L12 on chromosome Xq13.1-2
71E6	303	1767	NM_000291	Hs.78771	0	12	phosphoglycerate kinase 1 (PGK1), mRNA /cds=(
181D8	2104	3677	NM_018834	Hs.78825	0	4	matrin 3 (MATR3), mRNA /cds=(254,2800) /gb=NM
126G6	2498	2959	AL162049	Hs.78829	0	1	mRNA; cDNA DKFZp762E1712 (from clone DKFZp762E
41C3	1743	2340	M31932	Hs.78864	0	2	IgG low affinity Fc fragment receptor (FcRIIa) mRNA, c
166D11	1696	2156	M81601	Hs.78869	0	1	transcription elongation factor (SII) mRNA, complete
517B3	565	1392	D42039	Hs.78871	0	3	mRNA for KIAA0081 gene, partial cds /cds=(0,702) /gb=
180G11	59	517	NM_020548	Hs.78888	0	1	diazepam binding inhibitor (GABA receptor mod
99B7	2356	3329	U07802	Hs.78909	0	45	Tis11d gene, complete cds /cds=(291,1739) /gb=U07802
54C4	557	1101	U13045	Hs.78915	0	1	nuclear respiratory factor-2 subunit beta 1 mRNA, com
44A5	634	1128	U29607	Hs.78935	0	2	methionine aminopeptidase mRNA, complete cds /cds=(2
63A2	964	1050	X92106	Hs.78943	7.00E-31	1	bleomycin hydrolase /cds=(78,1445) /gb
163G9	228	877	L13463	Hs.78944	0	3	helix-loop-helix basic phosphoprotein (G0S8) mRNA,
119H6	472	877	NM_002923	Hs.78944	0	1	regulator of G-protein signalling 2, 24kD (RG
166E2	5629	5764	U51903	Hs.78993	2.00E-69	1	RasGAP-related protein (IQGAP2) mRNA, complete cds
40F9	66	603	M15796	Hs.78996	0	1	cyclin protein gene, complete cds /cds=(118,903) /gb
593E5	156	854	NM_012245	Hs.79008	0	5	SKI-INTERACTING PROTEIN (SNW1), mRNA /cds=(2
485B7	276	599	AF063591	Hs.79015	1.00E-136	1	brain my033 protein mRNA, complete cds /cds=(5
61B4	125	732	X05323	Hs.79015	0	2	MRC OX-2 gene signal sequence /cds=(0,824) /gb=X05323
71C8	330	1958	NM_005261	Hs.79022	0	24	GTP-binding protein overexpressed in skeletal
75G8	330	1957	U10550	Hs.79022	0	63	Gem GTPase (gem) mRNA, complete cds /cds=(213,1103) /
584G1	4424	5153	AF226044	Hs.79025	0	2	HSNFRK (HSNFRK) mRNA, complete cds /cds=(641,2
117C5	358	933	NM_012413	Hs.79033	0	1	glutaminyl-peptide cyclotransferase (glutam
72B2	910	2015	AJ250915	Hs.79037	0	9	p10 gene for chaperonin 10 (Hsp10 protein) and
71G11	880	1981	NM_002156	Hs.79037	0	5	heat shock 60kD protein 1 (chaperonin) (HSPD1)

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

193H12	1859	2474	NM_003243	Hs.79059	0	5	transforming growth factor, beta receptor III
460B4	846	1325	NM_001930	Hs.79064	0	1	deoxyhypusine synthase (DHPS), transcript va
75C4	1166	2087	K02276	Hs.79070	0	85	(Daudi) translocated t(8;14) c-myc oncogene mRNA, co
71G10	1274	2121	NM_002467	Hs.79070	0	12	v-myc avian myelocytomatosis viral oncogene h
183D8	385	741	NM_002710	Hs.79081	0	1	protein phosphatase 1, catalytic subunit, gam
170A12	741	1203	X74008	Hs.79081	0	1	protein phosphatase 1 gamma /cds=(154,11
121D9	2920	3385	NM_006378	Hs.79089	0	1	sema domain, immunoglobulin domain (Ig), tran
40C12	2933	4108	U60800	Hs.79089	0	4	semaphorin (CD100) mRNA, complete cds /cds=(87,2675)
104E1	1708	1932	L35263	Hs.79107	1.00E-101	1	CSaids binding protein (CSBP1) mRNA, complete cds /cd
70B2	913	2497	AK000221	Hs.79110	0	9	FLJ20214 fis, clone COLF2014, highly simi
123B12	1929	2644	D42043	Hs.79123	0	3	mRNA for KIAA0084 gene, partial cds /cds=(0,1946) /gb
193G7	802	1425	NM_004379	Hs.79194	0	2	cAMP responsive element binding protein 1 (CR
75D5	158	2139	NM_004233	Hs.79197	0	16	CD83 antigen (activated B lymphocytes, immuno
74H2	98	1357	NM_001154	Hs.79274	0	2	annexin A5 (ANXA5), mRNA /cds=(192,1154) /gb=
519G7	5358	5496	D86985	Hs.79276	2.00E-69	1	mRNA for KIAA0232 protein, partial cds /cds=(0,
462C2	1477	2031	NM_003006	Hs.79283	0	1	selectin P ligand (SELPLG), mRNA /cds=(59,1267
65C6	23	1609	M15353	Hs.79306	0	6	cap-binding protein mRNA, complete cds /cds=(1
64H8	326	1610	NM_001968	Hs.79306	0	3	eukaryotic translation initiation factor 4E
52C3	1333	1904	X64318	Hs.79334	0	1	E4BP4 gene /cds=(213,1601) /gb=X64318 /gi=30955
39F7	1179	1740	AF109733	Hs.79335	0	1	SWI/SNF-related, matrix-associated, actin-d
194A7	1512	1803	NM_003076	Hs.79335	1.00E-118	1	SWI/SNF related, matrix associated, actin dep
463E12	4326	4831	NM_015148	Hs.79337	0	1	KIAA0135 protein (KIAA0135), mRNA /cds=(1803,
526B5	1420	1867	NM_002958	Hs.79350	0	2	RYK receptor-like tyrosine kinase (RYK), mRNA
460F3	1755	2242	NM_006285	Hs.79358	0	2	testis-specific kinase 1 (TESK1), mRNA /cds=(
98B11	2076	4834	X76061	Hs.79362	0	11	H.sapiens p130 mRNA for 130K protein /cds=(69,3488) /gb=X76
45F3	2286	2666	NM_001423	Hs.79368	0	1	epithelial membrane protein 1 (EMP1), mRNA /cd
50C10	2016	2666	Y07909	Hs.79368	0	2	Progression Associated Protein /cds=(21
118E3	549	1078	NM_012198	Hs.79381	0	1	grancalcin (GCL), mRNA /cds=(119,772) /gb=NM_
181F4	657	1271	NM_002805	Hs.79387	0	2	proteasome (prosome, macropain) 26S subunit,
105H3	1114	1538	D83018	Hs.79389	0	1	for nel-related protein 2, complete cds /
173B2	429	3009	NM_006159	Hs.79389	0	5	nel (chicken)-like 2 (NELL2), mRNA /cds=(96,25
177B3	662	991	AC004382	Hs.79402	0	1	Chromosome 16 BAC clone CIT987SK-A-152E5 /cds
590H3	663	1002	NM_002694	Hs.79402	0	1	polymerase (RNA) II (DNA directed) polypeptide
523B7	223	582	NM_002946	Hs.79411	0	1	replication protein A2 (32kD) (RPA2), mRNA /c
182B10	472	1024	U02019	Hs.79625	1.00E-121	2	AU-rich element RNA-binding protein AUF1 mRNA, comple
479F3	100	301	NM_001783	Hs.79630	2.00E-86	1	CD79A antigen (immunoglobulin-associated al
40H9	582	1107	U05259	Hs.79630	0	1	MB-1 gene, complete cds /cds=(36,716) /gb=U05259 /gi
116A2	1003	1368	NM_006224	Hs.79709	1.00E-176	1	phosphotidylinositol transfer protein (PITPN
74G8	252	1297	D21853	Hs.79768	0	5	KIAA0111 gene, complete cds /cds=(214,1449)
525G2	830	1297	NM_014740	Hs.79768	0	2	KIAA0111 gene product (KIAA0111), mRNA /cds=(
125G3	2757	3339	AF072928	Hs.79877	0	1	myotubularin related protein 6 mRNA, partial c
184A2	532	1102	AF135162	Hs.79933	0	1	cyclin I (CYC1) mRNA, complete cds /cds=(199,13
514C6	329	1256	NM_006835	Hs.79933	0	6	cyclin I (CCNI), mRNA /cds=(0,1133) /gb=NM_006
116G5	824	1058	NM_006875	Hs.80205	1.00E-121	1	pim-2 oncogene (PIM2), mRNA /cds=(185,1189) /
106C11	1700	1995	U77735	Hs.80205	1.00E-125	1	pim-2 protooncogene homolog pim-2h mRNA, complete cd
110E3	276	653	AL136139	Hs.80261	0	1	DNA sequence from clone RP4-761I2 on chromosome 6 Con
478D1	1067	2761	NM_006403	Hs.80261	2.00E-70	2	enhancer of filamentation 1 (cas-like docking;

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

178C8	880	1226	AL050192	Hs.80285	0	1	mRNA; cDNA DKFZp586C1723 (from clone DKFZp586C
494F11	477	5535	NM_014739	Hs.80338	0	8	KIAA0164 gene product (KIAA0164), mRNA /cds=(
190A1	1165	1540	NM_004156	Hs.80350	1.00E-166	2	protein phosphatase 2 (formerly 2A), catalytic
461A1	4639	4913	NM_004653	Hs.80358	1.00E-140	1	SMC (mouse) homolog, Y chromosome (SMCY), mRNA
158A8	2656	3229	L24498	Hs.80409	0	1	gadd45 gene, complete cds /cds=(2327,2824) /gb=L2449
41E6	2385	2992	U84487	Hs.80420	0	2	CX3C chemokine precursor, mRNA, alternatively splice
40H4	2830	3605	NM_000129	Hs.80424	0	1	coagulation factor XIII, A1 polypeptide (F13A
464D3	214	835	NM_004899	Hs.80426	0	2	brain and reproductive organ-expressed (TNFR
75H8	1180	4930	U12767	Hs.80561	0	60	mitogen induced nuclear orphan receptor (MINOR) mRNA
593E10	1	510	NM_004552	Hs.80595	1.00E-158	5	NADH dehydrogenase (ubiquinone) Fe-S protein
113C5	1182	1583	NM_003336	Hs.80612	0	1	ubiquitin-conjugating enzyme E2A (RAD6 homol
515B7	268	538	NM_001020	Hs.80617	2.00E-91	3	ribosomal protein S16 (RPS16), mRNA /cds=(37,4
477F12	460	606	NM_018996	Hs.80618	1.00E-47	1	hypothetical protein (FLJ20015), mRNA /cds=(
41A8	1331	1788	L78440	Hs.80642	0	1	STAT4 mRNA, complete cds /cds=(81,2327) /gb=L
594C1	1594	2586	NM_003151	Hs.80642	0	4	signal transducer and activator of transcripti
112C8	1802	1932	NM_002198	Hs.80645	2.00E-35	1	interferon regulatory factor 1 (IRF1), mRNA /
522H8	1130	1533	NM_003355	Hs.80658	1.00E-135	4	uncoupling protein 2 (mitochondrial, proton c
123E4	259	757	NM_002129	Hs.80684	0	4	high-mobility group (nonhistone chromosomal)
109H1	263	754	X62534	Hs.80684	0	1	HMG-2 mRNA /cds=(214,843) /gb=X62534 /gi=32332
149G9	1020	1607	J05032	Hs.80758	0	2	aspartyl-tRNA synthetase alpha-2 subunit mRNA, compl
461F12	1702	2246	AL031600	Hs.80768	0	1	DNA sequence from clone 390E6 on chromosome 16. Contai
102B2	1486	2008	M16038	Hs.80887	0	1	lyn mRNA encoding a tyrosine kinase /cds=(297,1835) /
125B11	1260	2013	NM_002350	Hs.80887	0	5	v-yes-1 Yamaguchi sarcoma viral related oncog
37C9	2901	5260	D79990	Hs.80905	0	8	KIAA0168 gene, complete cds /cds=(196,1176)
196D6	2949	5261	NM_014737	Hs.80905	0	9	Ras association (RalGDS/AF-6) domain family 2
584H1	4072	4296	NM_002693	Hs.80961	3.00E-91	1	polymerase (DNA directed), gamma (POLG), nucl
584F9	31	568	AF174605	Hs.81001	0	5	F-box protein Fbx25 (FBX25) mRNA, partial cds
102D11	1037	1632	J03459	Hs.81118	0	1	leukotriene A-4 hydrolase mRNA, complete cds /cds=(68
193F8	1037	1643	NM_000895	Hs.81118	0	2	leukotriene A4 hydrolase (LTA4H), mRNA /cds=(
118H7	354	1148	U65590	Hs.81134	0	5	IL-1 receptor antagonist IL-1Ra (IL-1RN) gene
41H1	2549	2936	X60992	Hs.81226	0	1	CD6 mRNA for T cell glycoprotein CD6 /cds=(120,152
171B9	2070	2479	AF248648	Hs.81248	0	1	RNA-binding protein BRUNOL2 (BRUNOL2) mRNA, c
590A6	291	512	NM_002961	Hs.81256	3.00E-66	1	S100 calcium-binding protein A4 (calcium prot
73H2	389	1481	M69043	Hs.81328	0	14	MAD-3 mRNA encoding IkB-like activity, complet
513G1	637	1481	NM_020529	Hs.81328	0	13	nuclear factor of kappa light polypeptide gene
488F2	1065	1417	NM_004499	Hs.81361	1.00E-180	4	heterogeneous nuclear ribonucleoprotein A/B
151C8	1260	1423	U76713	Hs.81361	1.00E-61	1	apobec-1 binding protein 1 mRNA, complete cds /cds=(15
593B9	41	954	NM_001688	Hs.81634	0	3	ATP synthase, H+ transporting, mitochondrial
104H12	352	912	X60221	Hs.81634	0	1	H+-ATP synthase subunit b /cds=(32,802)
141G8	1132	1642	AK001883	Hs.81648	0	1	FLJ11021 fis, clone PLACE1003704, weakly
41A1	4214	4395	X06182	Hs.81665	5.00E-67	1	c-kit proto-oncogene mRNA /cds=(21,2951) /gb=X06182
102F5	3037	3646	D38551	Hs.81848	0	1	KIAA0078 gene, complete cds /cds=(184,2079)
111E11	1375	1752	NM_006265	Hs.81848	0	1	RAD21 (S. pombe) homolog (RAD21), mRNA /cds=(1
592F8	38	720	NM_014736	Hs.81892	0	1	KIAA0101 gene product (KIAA0101), mRNA /cds=(

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

194F1	6886	7115	AF241785	Hs.81897	1.00E-117	1	NPD012 (NPD012) mRNA, complete cds /cds=(552,2
525C6	1	615	NM_005563	Hs.81915	0	4	leukemia-associated phosphoprotein p18 (sta
101D12	3249	3508	D38555	Hs.81964	1.00E-143	1	KIAA0079 gene, complete cds /cds=(114,3491)
176D11	2996	3168	NM_004922	Hs.81964	9.00E-94	2	SEC24 (S. cerevisiae) related gene family, mem
129B7	5068	5759	D50683	Hs.82028	0	4	for TGF-betaIIR alpha, complete cds /cds=
195H6	946	1208	NM_006023	Hs.82043	6.00E-74	1	D123 gene product (D123), mRNA /cds=(280,1290)
481D9	2709	3085	NM_002184	Hs.82065	1.00E-134	1	interleukin 6 signal transducer (gp130, oncos
129A5	1338	1802	M14083	Hs.82085	0	1	beta-migrating plasminogen activator inhibitor I mR
57G9	500	1561	AF220656	Hs.82101	1.00E-145	3	apoptosis-associated nuclear protein PHLDA1
40C11	3748	4497	M27492	Hs.82112	0	1	interleukin 1 receptor mRNA, complete cds /cds=(82,17
481B6	3164	3609	NM_000877	Hs.82112	0	1	interleukin 1 receptor, type I (IL1R1), mRNA /
40H6	161	557	AB049113	Hs.82113	0	1	DUT mRNA for dUTP pyrophosphatase, complete cd
592B7	184	568	NM_001948	Hs.82113	1.00E-111	2	dUTP pyrophosphatase (DUT), mRNA /cds=(29,523
114F1	465	720	U70451	Hs.82116	1.00E-135	1	myeloid differentiation primary response protein My
71H5	194	3415	NM_006186	Hs.82120	0	36	nuclear receptor subfamily 4, group A, member
75C1	1264	3422	X75918	Hs.82120	0	84	NOT /cds=(317,2113) /gb=X75918 /gi=4158
40D1	1621	2080	M90391	Hs.82127	0	1	putative IL-16 protein precursor, mRNA, comple
71C4	678	5065	NM_002460	Hs.82132	0	88	interferon regulatory factor 4 (IRF4), mRNA /
75G12	3219	5316	U52682	Hs.82132	0	27	lymphocyte specific interferon regulatory factor/in
193G6	1118	2682	NM_006874	Hs.82143	1.00E-178	3	E74-like factor 2 (ets domain transcription fa
147F6	1484	1951	AK025643	Hs.82148	0	1	FLJ21990 fis, clone HEP06386 /cds=(22,49
155E4	853	1264	M64992	Hs.82159	0	1	prosomal protein P30-33K (pros-30) mRNA, complete cd
595F1	30	614	NM_002786	Hs.82159	0	3	proteasome (prosome, macropain) subunit, alp
58A4	473	1715	NM_005655	Hs.82173	0	3	TGFB inducible early growth response (TIEG), m
67E6	784	2109	S81439	Hs.82173	0	7	EGR alpha=early growth response gene alpha [human, prostate
593H2	132	722	NM_000985	Hs.82202	0	2	ribosomal protein L17 (RPL17), mRNA /cds=(138,
40H5	283	1442	M37033	Hs.82212	0	12	CD53 glycoprotein mRNA, complete cds /cds=(93,752) /
592C4	1	1442	NM_000560	Hs.82212	0	11	CD53 antigen (CD53), mRNA /cds=(93,752) /gb=N
460D4	1519	1845	NM_002510	Hs.82226	1.00E-160	1	glycoprotein (transmembrane) nmb (GPNMB), mR
61A8	507	736	AF045229	Hs.82280	1.00E-116	1	regulator of G protein signaling 10 mRNA, compl
45F7	418	651	NM_002925	Hs.82280	1.00E-119	1	regulator of G-protein signalling 10 (RGS10),
49C2	416	1323	NM_006417	Hs.82316	0	7	interferon-induced, hepatitis C-associated
41C11	847	1716	X63717	Hs.82359	0	2	APO-1 cell surface antigen /cds=(220,122
71H4	15	1627	NM_001781	Hs.82401	0	21	CD69 antigen (p60, early T-cell activation ant
75B10	9	1627	Z22576	Hs.82401	0	33	CD69 gene /cds=(81,680) /gb=Z22576 /gi=397938 /
117B7	1441	1515	NM_022059	Hs.82407	7.00E-28	1	CXC chemokine ligand 16 (CXCL16), mRNA /cds=(4
110D6	1219	1721	AF006088	Hs.82425	0	1	Arp2/3 protein complex subunit p16-Arc (ARC16)
598F10	39	1497	NM_005717	Hs.82425	0	5	actin related protein 2/3 complex, subunit 5 (
99A9	621	1214	D26018	Hs.82502	0	1	mRNA for KIAA0039 gene, partial cds /cds=(0,1475) /gb
183F6	222	2235	NM_001637	Hs.82542	0	2	acyloxyacyl hydrolase (neutrophil) (AOAH), m
459G4	5196	5801	NM_003682	Hs.82548	0	1	MAP-kinase activating death domain (MADD), mR
75A6	301	2231	D85429	Hs.82646	0	44	heat shock protein 40, complete cds /c
64A5	300	2008	NM_006145	Hs.82646	0	17	heat shock 40kD protein 1 (HSPF1), mRNA /cds=(4
50E5	628	2399	AK025459	Hs.82689	0	2	FLJ21806 fis, clone HEP00829, highly sim
115C6	23	589	NM_005087	Hs.82712	0	1	fragile X mental retardation, autosomal homol
105H10	1017	1429	M61199	Hs.82767	0	1	cleavage signal 1 protein mRNA, complete cds /cds=(97,
461A11	204	748	NM_006296	Hs.82771	0	1	vaccinia related kinase 2 (VRK2), mRNA /cds=(1

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

39B4	1049	1203	M25393	Hs.82829	8.00E-83	1	protein tyrosine phosphatase (PTPase) mRNA, complete
590F5	123	436	NM_002828	Hs.82829	1.00E-178	1	protein tyrosine phosphatase, non-receptor t
517F10	1038	2618	AK025583	Hs.82845	0	9	cDNA: FLJ21930 fis, clone HEP04301, highly sim
40B7	972	1933	M25280	Hs.82848	0	6	lymph node homing receptor mRNA, complete cds /cds=(11
515B1	1	2322	NM_000655	Hs.82848	0	12	selectin L (lymphocyte adhesion molecule 1) (
587A10	190	685	NM_001344	Hs.82890	0	1	defender against cell death 1 (DAD1), mRNA /cd
113G9	1	2812	AF208850	Hs.82911	0	7	BM-008 mRNA, complete cds /cds=(341,844) /gb=
127H6	1828	2501	NM_003591	Hs.82919	0	2	cullin 2 (CUL2), mRNA /cds=(146,2383) /gb=NM_0
477E3	931	1777	NM_006416	Hs.82921	0	2	solute carrier family 35 (CMP-sialic acid tran
184D2	1355	1773	AL049795	Hs.83004	1.00E-164	1	DNA sequence from clone RP4-622L5 on chromosome 1p34.
41F10	507	774	D49950	Hs.83077	1.00E-150	1	for interferon-gamma inducing factor(IGI
482E7	499	774	NM_001562	Hs.83077	5.00E-97	1	interleukin 18 (interferon-gamma-inducing f
515C6	111	1162	L38935	Hs.83086	1.00E-107	2	GT212 mRNA /cds=UNKNOWN /gb=L38935 /gi=100884
479D3	1775	2028	NM_001760	Hs.83173	1.00E-122	1	cyclin D3 (CCND3), mRNA /cds=(165,1043) /gb=N
583H12	945	1655	NM_012151	Hs.83363	0	9	coagulation factor VIII-associated (intronic
47B3	2140	3625	M58603	Hs.83428	0	13	nuclear factor kappa-B DNA binding subunit (NF-kappa-
58G1	2538	3625	NM_003998	Hs.83428	0	4	nuclear factor of kappa light polypeptide gene
477C6	1628	2131	Z49995	Hs.83465	0	1	H.sapiens mRNA (non-coding; clone h2A) /cds=UNKNOWN /gb=Z4
587D10	1576	1900	AF064839	Hs.83530	0	2	map 3p21; 3.15 cR from WI-9324 region, complete
516B9	1662	3296	X59405	Hs.83532	0	4	H.sapiens, gene for Membrane cofactor protein /cds=UNKNOWN
459A5	120	298	NM_017459	Hs.83551	7.00E-42	1	microfibrillar-associated protein 2 (MFAP2),
591A12	321	1116	NM_005731	Hs.83583	0	17	actin related protein 2/3 complex, subunit 2 (
102C1	554	1127	AK025198	Hs.83623	0	1	FLJ21545 fis, clone COL06195 /cds=UNKNOW
458C8	1022	1831	NM_001619	Hs.83636	0	1	adrenergic, beta, receptor kinase 1 (ADRBK1),
107G1	303	1008	L20688	Hs.83656	0	4	GDP-dissociation inhibitor protein (Ly-GDI) mRNA, c
597F8	293	1180	NM_001175	Hs.83656	0	55	Rho GDP dissociation inhibitor (GDI) beta (AR
591G5	1	216	NM_003142	Hs.83715	1.00E-108	3	Sjogren syndrome antigen B (autoantigen La) (
184H9	240	392	X69804	Hs.83715	4.00E-77	2	for La/SS-B protein /cds=UNKNOWN /gb=X69804
193C10	1	1605	BC000957	Hs.83724	1.00E-154	4	Similar to hypothetical protein MNCb-2146, c
40A2	1101	1294	U90904	Hs.83724	1.00E-72	1	clone 23773 mRNA sequence /cds=UNKNOWN /gb=U90904 /g
57H2	191	422	NM_001827	Hs.83758	1.00E-126	1	CDC28 protein kinase 2 (CKS2), mRNA /cds=(95,33
60E10	191	422	X54942	Hs.83758	1.00E-129	1	ckshs2 mRNA for Cks1 protein homologue /cds=(95,3
164F5	1896	2293	NM_016325	Hs.83761	0	1	zinc finger protein 274 (ZNF274), mRNA /cds=(4
463E6	555	1128	NM_000791	Hs.83765	0	1	dihydrofolate reductase (DHFR), mRNA /cds=(47
194F8	1806	2223	NM_002199	Hs.83795	1.00E-161	1	interferon regulatory factor 2 (IRF2), mRNA /
520D11	180	1229	NM_000365	Hs.83848	0	5	triosephosphate isomerase 1 (TPI1), mRNA /cds
168B6	530	891	U47924	Hs.83848	0	1	chromosome 12p13 sequence /cds=(373,1122) /gb=U4792
331E11	2591	3485	NM_000480	Hs.83918	0	8	adenosine monophosphate deaminase (isoform E
458A11	125	409	NM_000396	Hs.83942	1.00E-108	1	cathepsin K (pycnodysostosis) (CTSK), mRNA /
185H2	2501	2690	NM_000195	Hs.83951	3.00E-85	1	Hermansky-Pudlak syndrome (HPS), mRNA /cds=(2
99D2	977	1191	NM_019006	Hs.83954	1.00E-97	1	protein associated with PRK1 (AWP1), mRNA /cds
167D5	2275	2755	NM_000211	Hs.83968	0	4	integrin, beta 2 (antigen CD18 (p95), lymphocyt
524B2	262	575	BF028896	Hs.83992	1.00E-155	1	601765270F1 cDNA, 5' end /clone=IMAGE:3997576
523B2	688	1065	NM_015937	Hs.84038	0	1	CGI-06 protein (LOC51604), mRNA /cds=(6,1730)
102F1	951	1416	M63180	Hs.84131	0	1	threonyl-tRNA synthetase mRNA, complete cds /cds=(13
589D5	863	1700	NM_006400	Hs.84153	0	3	dynactin 2 (p50) (DCTN2), mRNA /cds=(136,1356)

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

108F6	448	704	U70439	Hs.84264	1.00E-117	1	silver-stainable protein SSP29 mRNA, complete cds /
146D6	1022	1253	K01144	Hs.84298	6.00E-95	2	major histocompatibility class II antigen gamma chain
188B10	823	1302	NM_004355	Hs.84298	0	1	CD74 antigen (invariant polypeptide of major
175D2	1060	1479	M63488	Hs.84318	1.00E-158	1	replication protein A 70kDa subunit mRNA complete cds
115F4	2305	2393	NM_002945	Hs.84318	2.00E-43	1	replication protein A1 (70kD) (RPA1), mRNA /cd
595H4	5400	5649	NM_004239	Hs.85092	1.00E-131	1	thyroid hormone receptor interactor 11 (TRIP1
106F1	493	1371	NM_017491	Hs.85100	0	3	WD repeat domain 1 (WDR1), transcript variant 1
40C10	438	880	X57025	Hs.85112	0	1	IGF-I mRNA for insulin-like growth factor I /cds=(166,
44C5	2247	2430	AF017257	Hs.85146	5.00E-89	1	chromosome 21 derived BAC containing erythrobl
45D4	1962	3324	X79067	Hs.85155	0	6	H.sapiens ERF-1 mRNA 3' end /cds=UNKNOWN /gb=X79067 /gi=483
591B9	2378	2603	NM_002880	Hs.85181	1.00E-109	1	v-raf-1 murine leukemia viral oncogene homolo
39E2	67	2493	X76488	Hs.85226	0	3	lysosomal acid lipase /cds=(145,1344) /
62H12	1249	1975	M12824	Hs.85258	0	3	T-cell differentiation antigen Leu-2/T8 mRNA, partia
40C8	4505	4856	X53587	Hs.85266	0	1	integrin beta 4 /cds=UNKNOWN /gb=X53587 /gi=
40E11	1983	2633	S53911	Hs.85289	0	1	CD34=glycoprotein expressed in lymphohematopoietic proge
135A2	121	695	BC001646	Hs.85301	0	2	clone MGC:2392, mRNA, complete cds /cds=(964,
459H4	33	244	AK027067	Hs.85567	2.00E-90	1	cDNA: FLJ23414 fis, clone HEP20704 /cds=(37,10
479A4	5556	5974	AB040974	Hs.85752	1.00E-171	1	mRNA for KIAA1541 protein, partial cds /cds=(9
146C3	1610	2062	AL049796	Hs.85769	0	1	DNA sequence from clone RP4-561L24 on chromosome 1p22
463H11	871	1153	NM_006546	Hs.86088	5.00E-83	1	IGF-II mRNA-binding protein 1 (IMP-1), mRNA /
480A12	2	165	NM_004876	Hs.86371	7.00E-84	1	zinc finger protein 254 (ZNF254), mRNA /cds=(1
192F7	2854	3462	AF198614	Hs.86386	0	3	Mcl-1 (MCL-1) and Mcl-1 delta S/TM (MCL-1) gene
459G3	12	577	AL049340	Hs.86405	0	1	mRNA; cDNA DKFZp564P056 (from clone DKFZp564P0
460E4	2361	2787	NM_000161	Hs.86724	0	2	GTP cyclohydrolase 1 (dopa-responsive dystoni
62F9	834	1282	M60724	Hs.86858	0	1	p70 ribosomal S6 kinase alpha-I mRNA, complete cds /cd
187E7	84	766	NM_001695	Hs.86905	0	1	ATPase, H+ transporting, lysosomal (vacuolar
159D4	315	559	J03798	Hs.86948	1.00E-113	1	autoantigen small nuclear ribonucleoprotein Sm-D mR
459F9	1557	1619	NM_006938	Hs.86948	2.00E-25	1	small nuclear ribonucleoprotein D1 polypeptid
480G11	87	603	BG168139	Hs.87113	0	1	602341526F1 cDNA, 5' end /clone=IMAGE:4449343
41D6	2208	2320	M35999	Hs.87149	4.00E-39	1	platelet glycoprotein IIIa (GPIIIa) mRNA, complete c
462H11	387	648	NM_003806	Hs.87247	1.00E-133	1	harakiri, BCL2-interacting protein (contains
99D7	614	5517	NM_003246	Hs.87409	0	62	thrombospondin 1 (THBS1), mRNA /cds=(111,3623
39B8	2130	5517	X14787	Hs.87409	0	33	thrombospondin /cds=(111,3623) /gb=X14787
525A2	329	560	NM_007047	Hs.87497	1.00E-129	2	butyrophilin, subfamily 3, member A2 (BTN3A2)
583F2	3303	3622	D63876	Hs.87726	1.00E-155	1	mRNA for KIAA0154 gene, partial cds /cds=(0,2080) /gb
184D7	2211	2556	M34181	Hs.87773	1.00E-165	1	testis-specific cAMP-dependent protein kinase catal
460A4	499	1074	AL117637	Hs.87794	0	1	mRNA; cDNA DKFZp434I225 (from clone DKFZp434I2
459G2	258	452	AW967701	Hs.87912	8.00E-88	1	EST379776 cDNA /gb=AW967701 /gi=8157540 /ug=
74H7	1660	2397	AK026960	Hs.88044	0	9	FLJ23307 fis, clone HEP11549, highly sim
463D12	351	568	AI184553	Hs.88130	1.00E-118	1	qd60a05.x1 cDNA, 3' end /clone=IMAGE:1733840
595B1	309	986	NM_003454	Hs.88219	0	1	zinc finger protein 200 (ZNF200), mRNA /cds=(2
458D3	1018	1285	NM_000487	Hs.88251	6.00E-74	1	arylsulfatase A (ARSA), mRNA /cds=(375,1898)
462F4	4272	4846	AJ271878	Hs.88414	0	1	mRNA for putative transcription factor (BACH2
460B12	1267	2022	NM_006800	Hs.88764	0	3	male-specific lethal-3 (Drosophila)-like 1

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

461A4	2039	2421	AL161659	Hs.88820	0	1	DNA sequence from clone RP11-526K24 on chromosome 20
460F9	3413	3654	NM_000397	Hs.88974	1.00E-133	1	cytochrome b-245, beta polypeptide (chronic g
459G9	790	1160	NM_006228	Hs.89040	1.00E-145	1	prepronociceptin (PNOC), mRNA /cds=(211,741)
70H12	1	661	AV716500	Hs.89104	0	274	AV716500 cDNA, 5' end /clone=DCBAKA08 /clone_
469H5	1620	2142	AB040961	Hs.89135	0	1	mRNA for KIAA1528 protein, partial cds /cds=(4
175G6	2069	2501	D83243	Hs.89385	0	1	NPAT mRNA, complete cds /cds=(66,4349) /gb=D83243 /g
592B10	3703	3936	NM_002519	Hs.89385	1.00E-130	1	nuclear protein, ataxia-telangiectasia locu
120B7	337	630	NM_005176	Hs.89399	1.00E-114	1	ATP synthase, H+ transporting, mitochondrial
39D2	370	1892	AF147204	Hs.89414	0	68	chemokine receptor CXCR4-Lo (CXCR4) mRNA, alt
99H4	7	1625	NM_003467	Hs.89414	0	137	chemokine (C-X-C motif), receptor 4 (fusin) (C
106D2	2	266	U03644	Hs.89421	1.00E-143	1	recepin mRNA, complete cds /cds=(32,1387) /gb=U03644
41F5	1203	1522	M16336	Hs.89476	1.00E-170	1	T-cell surface antigen CD2 (T11) mRNA, complete cds, c
463A3	876	1025	NM_000698	Hs.89499	1.00E-79	1	arachidonate 5-lipoxygenase (ALOX5), mRNA /c
47D12	1198	4887	AB028969	Hs.89519	0	2	for KIAA1046 protein, complete cds /cds=(
498G2	4420	5265	NM_014928	Hs.89519	0	2	KIAA1046 protein (KIAA1046), mRNA /cds=(577,1
589G3	598	689	NM_002796	Hs.89545	4.00E-45	2	proteasome (prosome, macropain) subunit, bet
331B1	699	788	S71381	Hs.89545	1.00E-41	1	prosome beta-subunit=multicatalytic proteinase complex
110A2	1403	1739	AK026432	Hs.89555	1.00E-177	1	FLJ22779 fis, clone KAIA1741 /cds=(234,1
118E4	780	1672	NM_002110	Hs.89555	0	5	hemopoietic cell kinase (HCK), mRNA /cds=(168,
41B8	570	1166	M89957	Hs.89575	0	1	immunoglobulin superfamily member B cell receptor co
44A11	2567	2808	L20814	Hs.89582	1.00E-115	1	glutamate receptor 2 (HBGR2) mRNA, complete cds /cds=(
191G11	309	596	NM_006284	Hs.89657	1.00E-162	11	TATA box binding protein (TBP)-associated fac
72G5	1172	1575	AX023367	Hs.89679	0	38	Sequence 38 from Patent WO0006605
71B12	40	559	NM_000586	Hs.89679	0	13	interleukin 2 (IL2), mRNA /cds=(47,517) /gb=N
179G12	158	737	M36821	Hs.89690	0	1	cytokine (GRO-gamma) mRNA, complete cds
193B5	680	1146	NM_002994	Hs.89714	0	17	small inducible cytokine subfamily B (Cys-X-Cy
182G10	681	1146	X78686	Hs.89714	0	7	ENA-78 mRNA /cds=(106,450) /gb=X78686 /gi=47124
191C6	617	1597	NM_021950	Hs.89751	0	2	membrane-spanning 4-domains, subfamily A, m
40H3	1347	1597	X07203	Hs.89751	3.00E-71	1	CD20 receptor (S7) /cds=(90,983) /gb=X07203
458H2	3524	4331	NM_002024	Hs.89764	0	2	fragile X mental retardation 1 (FMR1), mRNA /c
40F6	1665	2210	D38081	Hs.89887	0	1	thromboxane A2 receptor, complete cds /cds=(9
473E1	578	956	AL515381	Hs.89986	1.00E-172	1	AL515381 cDNA /clone=CL0BB017ZH06-(3-prime)
126A12	770	982	AL558028	Hs.90035	1.00E-102	1	AL558028 cDNA /clone=CS0DJ002YF02-(5-prime)
183E12	2203	2814	NM_001316	Hs.90073	0	1	chromosome segregation 1 (yeast homolog)-like
145H12	1602	1811	AK026766	Hs.90077	1.00E-113	2	FLJ23113 fis, clone LNG07875, highly sim
62C2	1472	2610	AB023420	Hs.90093	0	2	for heat shock protein apg-2, complete cds
46H6	3172	3411	D26488	Hs.90315	6.00E-86	1	mRNA for KIAA0007 gene, partial cds /cds=(0,2062) /gb
116E2	1637	2016	AK025800	Hs.90421	1.00E-118	1	cDNA: FLJ22147 fis, clone HEP22163, highly sim
525H3	6	1231	NM_004261	Hs.90606	0	2	15 kDa selenoprotein (SEP15), mRNA /cds=(4,492
184D8	287	387	BE888304	Hs.90654	1.00E-46	2	601514033F1 cDNA, 5' end /clone=IMAGE:3915177
99D4	1948	4309	D50918	Hs.90998	0	5	mRNA for KIAA0128 gene, partial cds /cds=(0,1276) /gb
72B9	571	1312	AK026954	Hs.91065	0	1	FLJ23301 fis, clone HEP11120 /cds=(2,188
586H8	189	478	NM_000987	Hs.91379	2.00E-83	1	ribosomal protein L26 (RPL26), mRNA /cds=(6,44
160A12	1	132	X69392	Hs.91379	4.00E-69	5	ribosomal protein L26 /cds=(6,443) /gb=
331H4	1632	2166	AK027210	Hs.91448	0	1	FLJ23557 fis, clone LNG09686, highly sim
473E6	915	1390	NM_004556	Hs.91640	0	2	nuclear factor of kappa light polypeptide gene
69E4	673	1328	AB007956	Hs.92381	1.00E-122	2	mRNA, chromosome 1 specific transcript KIAA04
182F10	117	781	AF070523	Hs.92384	0	1	JWA protein mRNA, complete cds /cds=(115,681)
585F10	77	1890	NM_006407	Hs.92384	0	13	vitamin A responsive; cytoskeleton related (J

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

469G3	2061	2293	AK025683	Hs.92414	1.00E-110	1	cDNA: FLJ22030 fis, clone HEP08669 /cds=UNKNOW
472H4	247	671	AW978555	Hs.92448	0	1	EST390664 cDNA /gb=AW978555 /gi=8169822 /ug=
193F11	2051	4721	NM_003103	Hs.92909	0	3	SON DNA binding protein (SON), mRNA /cds=(414,4
37E7	1287	1805	AK002059	Hs.92918	0	1	FLJ11197 fis, clone PLACE1007690 /cds=(37
111D7	244	596	NM_016623	Hs.92918	1.00E-166	1	hypothetical protein (BM-009), mRNA /cds=(385
41B10	1216	1530	U24577	Hs.93304	1.00E-173	1	LDL-phospholipase A2 mRNA, complete cds /cds=(216,15
48B4	76	723	NM_001417	Hs.93379	0	5	eukaryotic translation initiation factor 4B
39F8	76	876	X55733	Hs.93379	0	1	initiation factor 4B cDNA /cds=(0,1835) /gb=X557
471B10	660	886	NM_007020	Hs.93502	1.00E-125	1	U1-snRNP binding protein homolog (70kD) (U1SN
467A3	1189	1284	X91348	Hs.93522	3.00E-36	1	H.sapiens predicted non coding cDNA (DGCR5) /cds=UNKNOWN /
461B5	652	874	NM_003367	Hs.93649	1.00E-104	1	upstream transcription factor 2, c-fos intera
62B8	1386	1739	J05016	Hs.93659	1.00E-170	1	(clone pA3) protein disulfide isomerase related prote
461E7	1931	2086	NM_004911	Hs.93659	1.00E-65	1	protein disulfide isomerase related protein (
458G11	2423	3161	AB040959	Hs.93836	0	1	mRNA for KIAA1526 protein, partial cds /cds=(0
104E3	516	981	AK000967	Hs.93872	0	1	FLJ10105 fis, clone HEMBA1002542 /cds=UN
41B6	87	846	X04430	Hs.93913	0	2	IFN-beta 2a mRNA for interferon-beta-2 /cds=(86,724)
179H7	1610	1682	AF009746	Hs.94395	9.00E-34	1	peroxisomal membrane protein 69 (PMP69) mRNA,
470G3	74	493	NM_007221	Hs.94446	0	1	polyamine-modulated factor 1 (PMF1), mRNA /c
472A5	2325	2429	AK022267	Hs.94576	2.00E-48	1	cDNA FLJ12205 fis, clone MAMMA1000931 /cds=UNK
459C9	5356	6120	NM_006421	Hs.94631	0	3	brefeldin A-inhibited guanine nucleotide-exc
465F8	3580	4049	NM_015125	Hs.94970	0	1	KIAA0306 protein (KIAA0306), mRNA /cds=(0,436
57B9	4145	4379	NM_005109	Hs.95220	1.00E-126	1	oxidative-stress responsive 1 (OSR1), mRNA /c
160D6	30	480	X01451	Hs.95327	0	2	gene for 20K T3 glycoprotein (T3-delta-chain) of T-c
512G1	1	415	BF107010	Hs.95388	1.00E-175	2	601824367F1 cDNA, 5' end /clone=IMAGE:4043920
593E11	24	273	BG291649	Hs.95835	1.00E-79	10	602385778F1 cDNA, 5' end /clone=IMAGE:4514827
41H2	1011	1306	M28170	Hs.96023	1.00E-114	1	cell surface protein CD19 (CD19) gene, complete cds /c
149G8	213	435	BF222826	Hs.96487	1.00E-119	2	7q23f06.x1 /clone=IMAGE /gb=BF222826 /g
101G7	2266	3173	AL133227	Hs.96560	0	2	DNA sequence from clone RP11-394O2 on chromosome 20 C
103E6	2840	3451	BC000143	Hs.96560	0	1	Similar to hypothetical protein FLJ11656, cl
107G5	226	2349	BF673956	Hs.96566	7.00E-24	1	602137338F1 cDNA, 5' end /clone=IMAGE:4274048
461A12	3602	4135	AB014555	Hs.96731	0	2	mRNA for KIAA0655 protein, partial cds /cds=(0
595A8	82	1571	NM_000734	Hs.97087	1.00E-147	10	CD3Z antigen, zeta polypeptide (TiT3 complex)
479H8	883	1378	NM_014373	Hs.97101	0	3	putative G protein-coupled receptor (GPCR150)
466D12	2001	5732	NM_012072	Hs.97199	0	2	complement component C1q receptor (C1QR), mRN
194B3	1835	2898	NM_002990	Hs.97203	0	2	small inducible cytokine subfamily A (Cys-Cys)
109E9	2880	3536	AF083322	Hs.97437	0	1	centriole associated protein CEP110 mRNA, com
459H5	9	230	BF438062	Hs.97896	1.00E-116	1	7q66e08.x1 cDNA /clone=IMAGE /gb=BF438062 /g
473A4	871	1327	NM_007015	Hs.97932	0	1	chondromodulin I precursor (CHM-I), mRNA /cds
466E9	1408	1808	AL442083	Hs.98026	1.00E-172	2	mRNA; cDNA DKFZp547D144 (from clone DKFZp547D1
460E3	1290	1687	AF038564	Hs.98074	0	1	atrophin-1 interacting protein 4 (AIP4) mRNA,
462E6	103	642	NM_016440	Hs.98289	0	1	VRK3 for vaccinia related kinase 3 (LOC51231),
460B8	114	546	AA418743	Hs.98306	1.00E-178	1	zv98f06.s1 cDNA, 3' end /clone=IMAGE:767843 /
124A8	1	157	NM_019044	Hs.98324	2.00E-69	1	hypothetical protein (FLJ10996), mRNA /cds=(
71B10	79	520	AI761058	Hs.98531	1.00E-112	34	wi69b03.x1 cDNA, 3' end /clone=IMAGE:2398541
49F1	36	435	AA913840	Hs.98903	0	1	ol39d11.s1 cDNA, 3' end /clone=IMAGE:1525845

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

462F6	61	201	AC006276	Hs.99093	2.00E-74	1	chromosome 19, cosmid R28379 /cds=(0,633) /gb
473A2	47	475	BE326857	Hs.99237	0	1	hr65h06.x1 cDNA, 3' end /clone=IMAGE:3133403
599D8	1468	1748	NM_005825	Hs.99491	1.00E-132	1	RAS guanyl releasing protein 2 (calcium and DA
459F8	300	541	AW444899	Hs.99665	1.00E-123	1	UI-H-BI3-ajz-d-07-0-UI.s1 cDNA, 3' end /clon
163H9	8	141	AL049319	Hs.99821	2.00E-58	1	cDNA DKFZp564C046 (from clone DKFZp564C0
165H8	1176	1930	NM_015400	Hs.99843	0	2	DKFZP586N0721 protein (DKFZP586N0721), mRNA
188C9	543	998	NM_001436	Hs.99853	0	2	fibrillarin (FBL), mRNA /cds=(59,1024) /gb=N
37H2	759	2017	AC018755	Hs.99855	0	4	chromosome 19, BAC BC330783 (CIT-HSPC_470E3),
127H3	758	2183	NM_001462	Hs.99855	0	5	formyl peptide receptor-like 1 (FPRL1), mRNA
62F2	1	642	BF315159	Hs.99858	0	6	601899519F1 cDNA, 5' end /clone=IMAGE:4128749
599A7	26	838	NM_000972	Hs.99858	0	11	ribosomal protein L7a (RPL7A), mRNA /cds=(31,8
167B3	1994	2101	AB032251	Hs.99872	2.00E-37	1	BPTF mRNA for bromodomain PHD finger transcript
41G8	461	751	L08096	Hs.99899	1.00E-161	1	CD27 ligand mRNA, complete cds /cds=(150,731) /gb=L08
479C10	327	738	NM_001252	Hs.99899	0	1	tumor necrosis factor (ligand) superfamily, m
36D8	1180	2315	AL162047	Hs.99908	0	7	cDNA DKFZp762E1112 (from clone DKFZp762E
593E2	62	435	NM_000983	Hs.99914	1.00E-145	1	ribosomal protein L22 (RPL22), mRNA /cds=(51,4
478C8	48	311	NM_000023	Hs.99931	1.00E-112	1	sarcoglycan, alpha (50kD dystrophin-associat
61A1	827	1053	S62140	Hs.99969	1.00E-126	1	TLS=translocated in liposarcoma [human, mRNA, 1824 nt] /cd
40C7	971	1724	X69819	Hs.99995	0	1	ICAM-3 mRNA /cds=(8,1651) /gb=X69819 /gi=32627
116F8	109	376	NM_002964	Hs.100000	1.00E-123	5	S100 calcium-binding protein A8 (calgranulin
121F4	30	540	NM_001629	Hs.100194	1.00E-118	7	arachidonate 5-lipoxygenase-activating pro
46G10	5175	5624	NM_003605	Hs.100293	0	2	O-linked N-acetylglucosamine (GlcNAc) transf
49E4	1279	2585	NM_006773	Hs.100555	0	4	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide
61E1	1279	1767	X98743	Hs.100555	0	2	RNA helicase (Myc-regulated dead box pro
460A10	824	1321	NM_018099	Hs.100895	0	1	hypothetical protein FLJ10462 (FLJ10462), mR
458F1	1	303	R18757	Hs.100896	1.00E-157	1	yg17e04.r1 cDNA, 5' end /clone=IMAGE:32522 /c
64B8	2062	2711	AB007859	Hs.100955	0	1	mRNA for KIAA0399 protein, partial cds /cds=(0,
515H6	131	201	NM_001207	Hs.101025	6.00E-33	1	basic transcription factor 3 (BTF3), mRNA /cd
472H12	10	358	AW968686	Hs.101340	0	1	EST380762 cDNA /gb=AW968686 /gi=8158527 /ug=
99G6	2427	4860	AB002384	Hs.101359	0	9	mRNA for KIAA0386 gene, complete cds /cds=(177,3383)
62E12	193	573	AI936516	Hs.101370	1.00E-100	6	wd28h07.x1 cDNA, 3' end /clone=IMAGE:2329501
493B9	3	638	AL583391	Hs.101370	0	8	AL583391 cDNA /clone=CS0DL012YA12-(3-prime)
117D4	2812	2966	NM_006291	Hs.101382	7.00E-79	1	tumor necrosis factor, alpha-induced protein
462A9	382	620	BC000764	Hs.101514	1.00E-133	1	hypothetical protein FLJ10342, clone MGC:27
193G3	3368	3659	AL139349	Hs.102178	3.00E-88	1	DNA sequence from clone RP11-261P9 on chromosome 20.
62H6	3035	4257	AF193339	Hs.102506	0	5	eukaryotic translation initiation factor 2 a
46E2	3223	4023	NM_004836	Hs.102506	0	2	eukaryotic translation initiation factor 2-a
460C4	151	635	AW978361	Hs.102630	0	2	EST390470 cDNA /gb=AW978361 /gi=8169626 /ug=
58E4	1	321	BF970875	Hs.102647	1.00E-177	2	602271536F1 cDNA, 5' end /clone=IMAGE:4359609
189G9	5473	6137	NM_018489	Hs.102652	0	2	hypothetical protein ASH1 (ASH1), mRNA /cds=(
111H5	3043	3331	AK000354	Hs.102669	1.00E-125	1	cDNA FLJ20347 fis, clone HEP13790 /cds=(708,14
465B8	27	348	AI707589	Hs.102793	1.00E-164	1	as30b05.x1 cDNA, 3' end /clone=IMAGE:2318673
126G11	1069	1431	NM_016128	Hs.102950	0	2	coat protein gamma-cop (LOC51137), mRNA /cds=
165H5	326	564	BF698884	Hs.103180	4.00E-71	1	602126455F1 cDNA, 5' end /clone=IMAGE:4283340
108H6	2135	2505	AB023187	Hs.103329	1.00E-59	1	for KIAA0970 protein, complete cds /cds=(
521C9	1440	1962	AL136885	Hs.103378	0	2	mRNA; cDNA DKFZp434P116 (from clone DKFZp434P1

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

458C9	3876	4415	AF254411	Hs.103521	0	1	ser/arg-rich pre-mRNA splicing factor SR-A1 (
99F6	349	767	NM_018623	Hs.103657	0	5	hypothetical protein PRO2219 (PRO2219), mRNA
162G11	1745	2161	AF117829	Hs.103755	1.00E-151	1	8q21.3: RICK gene /cds=(224,1846) /gb=AF11782
188G1	1757	2566	NM_004501	Hs.103804	0	2	heterogeneous nuclear ribonucleoprotein U (
470F7	56	302	NM_024056	Hs.103834	1.00E-137	1	hypothetical protein MGC5576 (MGC5576), mRNA
460A11	225	288	BG033732	Hs.103902	3.00E-29	1	602301101F1 cDNA, 5' end /clone=IMAGE:4402465
522H7	2157	2397	NM_006342	Hs.104019	1.00E-132	1	transforming, acidic coiled-coil containing
39E5	1007	2535	L12168	Hs.104125	0	10	adenylyl cyclase-associated protein (CAP) mRN
98C11	1023	2558	NM_006367	Hs.104125	0	29	adenylyl cyclase-associated protein (CAP), m
461B2	88	221	AW968823	Hs.104157	1.00E-38	1	EST380899 cDNA /gb=AW968823 /gi=8158664 /ug=
110A4	4010	4306	AB023143	Hs.104305	1.00E-125	1	for KIAA0926 protein, complete cds /cds=(
122H5	4634	5232	NM_014922	Hs.104305	0	2	KIAA0926 protein (KIAA0926), mRNA /cds=(522,4
105C2	1817	2174	AB020669	Hs.104315	0	1	for KIAA0862 protein, complete cds /cds=(
37G4	1321	2886	AF016495	Hs.104624	0	46	small solute channel 1 (SSC1) mRNA, complete cd
98D4	1578	2946	NM_020980	Hs.104624	0	71	aquaporin 9 (AQP9), mRNA /cds=(286,1173) /gb=
458E6	1007	1399	NM_015898	Hs.104640	0	1	HIV-1 inducer of short transcripts binding pro
462C11	1037	1532	NM_018492	Hs.104741	0	1	PDZ-binding kinase; T-cell originated protein
118G4	1940	2513	BC002538	Hs.104879	0	2	serine (or cysteine) proteinase inhibitor, c
496A7	1	618	BG035120	Hs.104893	0	4	602324815F1 cDNA, 5' end /clone=IMAGE:4413099
112G4	3421	3933	NM_003633	Hs.104925	0	2	ectodermal-neural cortex (with BTB-like doma
460E2	16	460	AI479075	Hs.104985	0	1	tm30h01.x1 cDNA, 3' end /clone=IMAGE:2158129
461H4	1500	1781	NM_020979	Hs.105052	1.00E-148	1	adaptor protein with pleckstrin homology and
469C7	231	380	NM_018331	Hs.105216	1.00E-77	1	hypothetical protein FLJ11125 (FLJ11125), mR
461B6	84	489	AA489227	Hs.105230	0	1	aa57f07.s1 cDNA, 3' end /clone=IMAGE:825061 /
462D5	1735	2129	NM_015393	Hs.105460	0	1	DKFZP564O0823 protein (DKFZP564O0823), mRNA
465H7	1	624	NM_017780	Hs.105461	0	1	hypothetical protein FLJ20357 (FLJ20357), mR
471F3	819	1126	AY007243	Hs.105484	1.00E-160	1	regenerating gene type IV mRNA, complete cds /
473C1	42	479	AW970759	Hs.105621	0	1	EST382842 cDNA /gb=AW970759 /gi=8160604 /ug=
102A9	1	331	AK025947	Hs.105664	0	1	FLJ22294 fis, clone HRC04426 /cds=(240,6
465G9	193	524	AI475680	Hs.105676	0	1	tc93d12.x1 cDNA, 3' end /clone=IMAGE:2073719
469G2	1528	1625	AK022481	Hs.105779	8.00E-38	1	cDNA FLJ12419 fis, clone MAMMA1003047, highly
482A9	289	839	NM_012483	Hs.105806	0	3	granulysin (GNLY), transcript variant 519, m
595B11	918	1300	NM_002343	Hs.105938	0	1	lactotransferrin (LTF), mRNA /cds=(294,2429)
69B3	3649	4226	Y13247	Hs.106019	0	1	fb19 mRNA /cds=(539,3361) /gb=Y13247 /gi=2117
459E8	106	563	NM_013322	Hs.106260	0	1	sorting nexin 10 (SNX10), mRNA /cds=(128,733)
459E2	1939	2361	NM_003171	Hs.106469	0	1	suppressor of var1 (S.cerevisiae) 3-like 1 (S
98H12	658	1040	BC002748	Hs.106650	0	2	Similar to hypothetical protein FLJ20533, cl
594H5	1418	1501	NM_001568	Hs.106673	6.00E-36	1	eukaryotic translation initiation factor 3,
194H12	751	1233	NM_021626	Hs.106747	0	1	serine carboxypeptidase 1 precursor protein (
138G6	2749	3214	AF189723	Hs.106778	0	3	calcium transport ATPase ATP2C1 (ATP2C1A) mRN
56A5	1	1089	AL355722	Hs.106875	0	2	EST from clone 35214, full insert /cds=UNKNOWN
67H8	844	1102	X71490	Hs.106876	1.00E-103	1	vacuolar proton ATPase, subunit D /cds=(2
463G10	538	725	AF035306	Hs.106890	1.00E-102	1	clone 23771 mRNA sequence /cds=UNKNOWN /gb=AF
121H2	14	394	NM_016619	Hs.107139	0	1	hypothetical protein (LOC51316), mRNA /cds=(
185D12	118	884	NM_001564	Hs.107153	0	3	inhibitor of growth family, member 1-like (ING
186D6	1140	1507	NM_017892	Hs.107213	0	1	hypothetical protein FLJ20585 (FLJ20585), mR
462B10	192	541	AI707896	Hs.107369	1.00E-168	1	as34a10.x1 cDNA, 3' end /clone=IMAGE:2319066
59A10	1694	2335	AJ270952	Hs.107393	0	3	for putative membrane protein (GENX-3745
499G1	2987	4266	AL035683	Hs.107526	1.00E-104	2	DNA sequence from clone RP5-1063B2 on chromosome 20q1
466F11	327	493	AI391443	Hs.107622	9.00E-90	1	tf96e06.x1 cDNA, 3' end /clone=IMAGE:2107138
182F9	153	649	AF265439	Hs.107707	0	1	DC37 mRNA, complete cds /cds=(5,856) /gb=AF26
481F9	1216	1609	NM_016270	Hs.107740	0	2	Kruppel-like factor (LOC51713), mRNA /cds=(84

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

184H4	189	576	AF081282	Hs.107979	0	1	small membrane protein 1 (SMP1) mRNA, complete
103E11	1006	2137	NM_014313	Hs.107979	0	4	small membrane protein 1 (SMP1), mRNA /cds=(99,
596H7	1265	1771	NM_004078	Hs.108080	0	3	cysteine and glycine-rich protein 1 (CSRP1), m
46H8	777	914	AF070640	Hs.108112	2.00E-47	1	clone 24781 mRNA sequence /cds=UNKNOWN /gb=AF
53B4	1552	1967	U32986	Hs.108327	0	2	xeroderma pigmentosum group E UV-damaged DNA binding
124A10	1089	1733	AK001428	Hs.108332	0	3	cDNA FLJ10566 fis, clone NT2RP2002959, highly
127F8	428	746	AL136941	Hs.108338	0	1	mRNA; cDNA DKFZp586C1924 (from clone DKFZp586
191G10	518	883	AL136640	Hs.108548	0	2	mRNA; cDNA DKFZp564F163 (from clone DKFZp564F1
458G8	2374	5101	NM_016227	Hs.108636	0	2	membrane protein CH1 (CH1), mRNA /cds=(124,434
58F11	735	798	NM_006963	Hs.108642	2.00E-28	1	zinc finger protein 22 (KOX 15) (ZNF22), mRNA /
118B5	2715	2797	AK022874	Hs.108779	2.00E-38	1	cDNA FLJ12812 fis, clone NT2RP2002498 /cds=(3,
110H2	18	661	AF026292	Hs.108809	0	1	chaperonin containing t-complex polypeptide
181G4	1008	1142	NM_006429	Hs.108809	2.00E-71	1	chaperonin containing TCP1, subunit 7 (eta) (C
189F11	415	615	AK024569	Hs.108854	2.00E-79	1	cDNA: FLJ20916 fis, clone ADSE00738, highly s
596F8	5958	6097	AB011087	Hs.108945	8.00E-48	1	mRNA for KIAA0515 protein, partial cds /cds=(0,
157D8	399	830	NM_016145	Hs.108969	0	1	PTD008 protein (PTD008), /cds=(233,553)
175E7	712	1849	AL133111	Hs.109150	0	2	mRNA; cDNA DKFZp434H068 (from clone DKFZp434H0
514E1	66	613	NM_012417	Hs.109219	0	4	retinal degeneration B beta (RDGBB), mRNA /cd
106A4	1864	2220	AJ011895	Hs.109281	1.00E-111	1	for HIV-1, Nef-associated factor 1 alpha
169E1	938	1331	AK024297	Hs.109441	0	2	FLJ14235 fis, clone NT2RP4000167 /cds=(82
100B8	1	191	NM_012456	Hs.109571	3.00E-85	1	translocase of inner mitochondrial membrane 1
115B7	983	1193	NM_007074	Hs.109606	1.00E-116	1	coronin, actin-binding protein, 1A (CORO1A),
62H11	1	626	BF245892	Hs.109641	1.00E-154	10	601864070F1 cDNA, 5' end /clone=IMAGE:4082465
595B2	4976	5286	AB040884	Hs.109694	1.00E-142	1	mRNA for KIAA1451 protein, partial cds /cds=(0
75H11	227	482	BF244603	Hs.109697	1.00E-129	1	601862620F1 cDNA, 5' end /clone=IMAGE:4080412
118G3	219	392	NM_024292	Hs.109701	2.00E-66	1	ubiquitin-like 5 (UBL5), mRNA /cds=(65,286) /
105A5	3271	3532	AL117407	Hs.109727	1.00E-147	2	cDNA DKFZp434D2050 (from clone DKFZp434D
481B7	1101	1201	NM_006026	Hs.109804	9.00E-42	1	H1 histone family, member X (H1FX), mRNA /cds=(
476H12	1018	1429	NM_004310	Hs.109918	0	3	ras homolog gene family, member H (ARHH), mRNA
144C8	1252	1429	Z35227	Hs.109918	7.00E-92	1	TTF for small G protein /cds=(579,1154) /gb=
141E10	630	1269	AK001779	Hs.110445	0	4	FLJ10917 fis, clone OVARC1000321 /cds=(18
494D8	4102	4476	NM_014918	Hs.110488	0	1	KIAA0990 protein (KIAA0990), mRNA /cds=(494,2
47C3	2298	2431	D86974	Hs.110613	1.00E-60	1	KIAA0220 gene, partial cds /cds=(0,1661) /gb
194C10	1210	1704	AL157477	Hs.110702	0	1	mRNA; cDNA DKFZp761E212 (from clone DKFZp761E2
192F1	3254	3686	NM_015726	Hs.110707	1.00E-150	2	H326 (H326), mRNA /cds=(176,1969) /gb=NM_0157
595B8	1148	1414	NM_003472	Hs.110713	1.00E-147	1	DEK oncogene (DNA binding) (DEK), mRNA /cds=(3
459F3	3337	3915	NM_001046	Hs.110736	0	1	solute carrier family 12 (sodium/potassium/ch
195F5	1051	1482	AK025557	Hs.110771	0	2	cDNA: FLJ21904 fis, clone HEP03585 /cds=UNKNOW
53B10	163	742	NM_020150	Hs.110796	0	1	SAR1 protein (SAR1), mRNA /cds=(100,696) /gb=
164B11	122	932	NM_016039	Hs.110803	0	5	CGI-99 protein (LOC51637), mRNA /cds=(161,895
594H4	982	1454	AK026528	Hs.111222	6.00E-95	3	cDNA: FLJ22875 fis, clone KAT02879 /cds=(30,51
50A10	1688	2095	AF119897	Hs.111334	0	2	PRO2760 mRNA, complete cds /cds=UNKNOWN /gb=A
102H11	175	498	AI436587	Hs.111377	1.00E-148	1	ti03d11.x1 cDNA, 3' end /clone=IMAGE:2129397
109G11	1324	1388	AB016811	Hs.111554	2.00E-29	1	for ADP ribosylation factor-like protein,
144E10	77	304	BF219474	Hs.111611	1.00E-122	2	601884269F1 5' end /clone=IMAGE:4102769

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

583C9	4	272	NM_000988	Hs.111611	1.00E-148	10	ribosomal protein L27 (RPL27), mRNA /cds=(17,4
111F4	31	380	NM_014463	Hs.111632	0	1	Lsm3 protein (LSM3), mRNA /cds=(29,337) /gb=N
106E6	2646	2892	AL096723	Hs.111801	1.00E-135	1	cDNA DKFZp564H2023 (from clone DKFZp564H
169A2	773	1015	D14696	Hs.111894	1.00E-135	2	KIAA0108 gene, complete cds /cds=(146,847) /
182D6	264	748	NM_014713	Hs.111894	0	1	lysosomal-associated protein transmembrane
460D11	205	452	AI557431	Hs.111973	4.00E-60	1	PT2.1_7_C05.r cDNA, 3' end /clone_end=3' /gb=
121A7	355	589	NM_020382	Hs.111988	1.00E-128	1	PR/SET domain containing protein 07 (SET07), m
476C12	254	463	AA442585	Hs.112071	1.00E-111	1	zv57f09.r1 cDNA, 5' end /clone=IMAGE:757769 /
172E7	469	736	AF228422	Hs.112242	1.00E-143	1	normal mucosa of esophagus specific 1 (NMES1)
108E10	4800	4901	AF071076	Hs.112255	6.00E-48	1	cell-line HeLa Nup98-Nup96 precursor, mRNA, c
47G12	1	301	BF237710	Hs.112318	1.00E-165	5	601842210F1 cDNA, 5' end /clone=IMAGE:4079930
599G7	38	455	NM_019059	Hs.112318	0	32	6.2 kd protein (LOC54543), mRNA /cds=(93,260)
469F9	226	546	NM_002638	Hs.112341	1.00E-107	1	protease inhibitor 3, skin-derived (SKALP) (P
589G11	482	1336	AK026396	Hs.112497	0	2	cDNA: FLJ22743 fis, clone HUV00901 /cds=UNKNOW
464F10	1686	1917	NM_002978	Hs.112842	1.00E-119	1	sodium channel, nonvoltage-gated 1, delta (SC
54B11	1	423	BF025727	Hs.113029	0	26	601670406F1 cDNA, 5' end /clone=IMAGE:3953425
591C5	31	469	NM_001028	Hs.113029	0	10	ribosomal protein S25 (RPS25), mRNA /cds=(71,4
585F4	1882	3918	AK027136	Hs.113205	1.00E-130	3	cDNA: FLJ23483 fis, clone KAIA04052 /cds=UNKNO
61B12	1168	2386	AF105253	Hs.113368	0	5	neuroendocrine secretory protein 55 mRNA, com
163D9	3470	4109	Y08890	Hs.113503	0	1	mRNA for Ran_GTP binding protein 5
466C4	276	946	AL359916	Hs.113872	0	1	DNA sequence from clone RP11-550O8 on chromosome 20 C
592C12	2506	2696	AF323540	Hs.114309	2.00E-80	1	apolipoprotein L-I mRNA, splice variant B, co
476A11	121	528	AA702108	Hs.114931	0	1	zi85e01.s1 cDNA, 3' end /clone=IMAGE:447576 /
109F4	3123	3521	D30783	Hs.115263	0	1	for epiregulin, complete cds /cds=(166,67
123D1	3123	3526	NM_001432	Hs.115263	0	1	epiregulin (EREG), mRNA /cds=(166,675) /gb=N
465D7	1	175	BG288391	Hs.115467	1.00E-94	1	602388053F1 cDNA, 5' end /clone=IMAGE:4517076
74H9	346	602	AK027114	Hs.115659	1.00E-108	1	FLJ23461 fis, clone HSI07757 /cds=UNKNOW
585E4	384	1146	NM_024061	Hs.115659	0	3	hypothetical protein MGC5521 (MGC5521), mRNA
462C1	945	1222	NM_024036	Hs.115960	1.00E-152	1	hypothetical protein MGC3103 (MGC3103), mRNA
464E4	1276	1635	AK023633	Hs.116278	1.00E-138	1	cDNA FLJ13571 fis, clone PLACE1008405 /cds=UNK
43B10	1601	1798	AF283777	Hs.116481	9.00E-47	1	clone TCBAP0702 mRNA sequence /cds=UNKNOWN /g
465G1	374	654	NM_001782	Hs.116481	5.00E-85	2	CD72 antigen (CD72), mRNA /cds=(108,1187) /gb
51G8	29	203	BF341330	Hs.116567	6.00E-26	1	602013274F1 cDNA, 5' end /clone=IMAGE:4149066
40D10	2694	3430	X68742	Hs.116774	0	1	integrin, alpha subunit /cds=UNKNOWN /g
107D1	1778	1943	U71383	Hs.117005	1.00E-84	1	OB binding protein-2 (OB-BP2) mRNA, complete cds /cds
459D4	2882	3522	AK025364	Hs.117268	0	1	cDNA: FLJ21711 fis, clone COL10156 /cds=UNKNOW
473E8	2104	2233	AB029016	Hs.117333	2.00E-65	3	mRNA for KIAA1093 protein, partial cds /cds=(0
458E2	88	627	AI825645	Hs.117906	0	2	wb75b09.x1 cDNA, 3' end /clone=IMAGE:2311481
163A7	1160	1420	X53793	Hs.117950	1.00E-109	1	ADE2H1 mRNA showing homologies to SAICAR syntheta
123B8	18	740	NM_002799	Hs.118065	0	1	proteasome (prosome, macropain) subunit, bet
583G3	924	1199	AB011182	Hs.118087	1.00E-155	4	mRNA for KIAA0610 protein, partial cds /cds=(0,
127A1	263	557	NM_006441	Hs.118131	1.00E-141	1	5,10-methenyltetrahydrofolate synthetase (
459A10	188	817	AL522477	Hs.118142	0	1	AL522477 cDNA /clone=CS0DB008YK14-(3-prime)
584A10	8484	8875	NM_003316	Hs.118174	0	1	tetratricopeptide repeat domain 3 (TTC3), mRN
52D4	1287	1752	AK026486	Hs.118183	0	1	FLJ22833 fis, clone KAIA4266 /cds=(479,8
470B6	68	532	BF030930	Hs.118303	0	1	601558648F1 cDNA, 5' end /clone=IMAGE:3828706
41B3	5041	5669	M14648	Hs.118512	0	1	cell adhesion protein (vitronectin) receptor alpha s

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

125B8	999	1573	NM_003733	Hs.118633	0	1	2'-5'oligoadenylate synthetase-like (OASL),
459D3	3	427	AI052447	Hs.118659	0	1	oz07g04.x1 cDNA, 3' end /clone=IMAGE:1674678
112F11	191	387	NM_006923	Hs.118684	1.00E-103	1	stromal cell-derived factor 2 (SDF2), mRNA /c
129E4	1727	1891	AL050404	Hs.118695	2.00E-86	1	DNA sequence from clone 955M13 on chromosome 20. Conta
126H2	1512	2209	NM_000358	Hs.118787	0	2	transforming growth factor, beta-induced, 68
598D9	817	1106	NM_001155	Hs.118796	1.00E-108	1	annexin A6 (ANXA6), transcript variant 1, mRN
331E6	89	475	BE311727	Hs.118857	0	1	601143334F1 cDNA, 5' end /clone=IMAGE:3507009
521C1	700	1180	NM_006292	Hs.118910	0	2	tumor susceptibility gene 101 (TSG101), mRNA
139E8	463	1198	AJ012506	Hs.118958	0	1	activated in tumor suppression, clone TSA
69H2	578	1117	U05040	Hs.118962	0	1	FUSE binding protein mRNA, complete cds /cds=(26,1960
461F1	1241	1715	AK024119	Hs.118990	0	1	cDNA FLJ14057 fis, clone HEMBB1000337 /cds=UNK
481E1	1682	1969	NM_017544	Hs.119018	1.00E-129	1	transcription factor NRF (NRF), mRNA /cds=(653
479B4	45	203	AL109806	Hs.119057	5.00E-43	1	DNA sequence from clone RP5-1153D9 on chromosome 20 C
520F1	177	672	NM_012423	Hs.119122	1.00E-148	8	ribosomal protein L13a (RPL13A), mRNA /cds=(1
477E4	46	1565	AL109786	Hs.119155	0	3	mRNA full length insert cDNA clone EUROIMAGE 81
166F10	304	814	M37583	Hs.119192	0	3	histone (H2A.Z) mRNA, complete cds /cds=(106,492) /g
592E5	302	814	NM_002106	Hs.119192	0	7	H2A histone family, member Z (H2AFZ), mRNA /cd
54B1	47	1144	AJ400717	Hs.119252	0	9	TPT1 gene for translationally controlled tumo
594H9	609	1013	NM_000520	Hs.119403	0	1	hexosaminidase A (alpha polypeptide) (HEXA),
492D9	30	272	NM_001004	Hs.119500	1.00E-135	2	ribosomal protein, large P2 (RPLP2), mRNA /cd
59H8	14	1890	NM_016091	Hs.119503	0	12	HSPC025 (HSPC025), mRNA /cds=(33,1727) /gb=N
525E8	12	446	NM_006432	Hs.119529	0	2	epididymal secretory protein (19.5kD) (HE1),
166G7	1323	2293	M88108	Hs.119537	0	3	p62 mRNA, complete cds /cds=(106,1437) /gb=M88108 /g
112D10	1054	1722	NM_006559	Hs.119537	0	1	GAP-associated tyrosine phosphoprotein p62
158E9	847	1273	AL022326	Hs.119598	0	1	DNA sequence from clone 333H23 on chromosome 22q12.1-1
161H7	738	1272	NM_000967	Hs.119598	0	1	ribosomal protein L3 (RPL3), mRNA /cds=(6,1217
168F8	284	778	M34671	Hs.119663	0	1	lymphocytic antigen CD59/MEM43 mRNA, complete cds /c
585C9	285	783	NM_000611	Hs.119663	0	1	CD59 antigen p18-20 (antigen identified by mo
143G12	753	1329	AK023975	Hs.119908	0	4	FLJ13913 fis, clone Y79AA1000231, highly
55D12	1107	1365	NM_015934	Hs.119908	1.00E-119	1	nucleolar protein NOP5/NOP58 (NOP5/NOP58), m
467E7	37	419	AI492066	Hs.119923	0	1	tg12b03.x1 cDNA, 3' end /clone=IMAGE:2108525
462C10	2669	3025	NM_012318	Hs.120165	0	1	leucine zipper-EF-hand containing transmembr
473F11	396	1006	AK025068	Hs.120170	0	1	cDNA: FLJ21415 fis, clone COL04030 /cds=(138,7
98E11	211	458	AW081455	Hs.120219	1.00E-114	2	xc31c07.x1 cDNA, 3' end /clone=IMAGE:2585868
471C8	60	301	NM_014487	Hs.120766	1.00E-120	1	nucleolar cysteine-rich protein (HSA6591), m
134C4	284	529	AK000470	Hs.120769	9.00E-98	1	cDNA FLJ20463 fis, clone KAT06143 /cds=UNKNOWN
469C10	1	441	AA677952	Hs.120891	0	1	zi14a06.s1 cDNA, 3' end /clone=IMAGE:430738 /
60C9	1022	1615	AB011421	Hs.120996	0	1	for DRAK2, complete cds /cds=(261,1379) /
461A7	738	1274	NM_014205	Hs.121025	0	1	chromosome 11 open reading frame 5 (C11ORF5), m
104A4	557	1942	D89974	Hs.121102	0	4	for glycosylphosphatidyl inositol-ancho
196C9	557	1463	NM_004665	Hs.121102	0	9	vanin 2 (VNN2), mRNA /cds=(11,1573) /gb=NM_004
467F4	4	328	AW972196	Hs.121210	1.00E-162	1	EST384285 cDNA /gb=AW972196 /gi=8162042 /ug=
587A12	224	367	AW975541	Hs.121572	1.00E-62	1	EST387650 cDNA /gb=AW975541 /gi=8166755 /ug=
36G5	13	604	AL008729	Hs.121591	0	1	DNA sequence from PAC 257A7 on chromosome 6p24. Contai
464C1	120	413	AA772692	Hs.121709	1.00E-120	1	ai35b09.s1 cDNA, 3' end /clone=1358969 /clone

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

36E2	411	821	AK025556	Hs.121849	0	1	FLJ21903 fis, clone HEP03579 /cds=(84,46
196A6	411	1113	NM_022818	Hs.121849	0	1	Microtubule-associated proteins 1A and 1B, l
471G2	176	333	AW469546	Hs.122116	2.00E-64	1	hd19e09.x1 cDNA, 3' end /clone=IMAGE:2909992
462F5	218	611	BF677944	Hs.122406	1.00E-166	1	602084766F1 cDNA, 5' end /clone=IMAGE:4248905
465A6	376	478	AV762642	Hs.122431	2.00E-28	1	AV762642 cDNA, 5' end /clone=MDSEMB08 /clone_
467G10	603	803	AL040371	Hs.122487	9.00E-96	1	DKFZp434P0213_r1 cDNA, 5' end /clone=DKFZp434
465C12	66	260	AI804629	Hs.122848	3.00E-83	1	tc81g03.x1 cDNA, 3' end /clone=IMAGE:2072596
98H6	442	591	AI081246	Hs.122983	5.00E-78	1	oy67b06.x1 cDNA, 3' end /clone=IMAGE:1670867
52B4	123	236	BE676541	Hs.123254	8.00E-46	1	7f31g03.x1 cDNA, 3' end /clone=IMAGE:3296308
128C7	4875	5186	AB020631	Hs.123654	1.00E-131	1	mRNA for KIAA0824 protein, partial cds /cds=(0
184B5	594	1187	AL109865	Hs.124186	0	1	DNA sequence from clone GS1-120K12 on chromosome 1q25
106A6	1135	1456	AK026776	Hs.124292	9.00E-99	1	FLJ23123 fis, clone LNG08039 /cds=UNKNOW
525G12	314	503	BF996704	Hs.124344	1.00E-72	1	MR1-GN0173-071100-009-g10 cDNA /gb=BF996704
466C3	120	496	AA831838	Hs.124391	1.00E-172	1	oc85h06.s1 cDNA, 3' end /clone=IMAGE:1356539
48G4	1	568	AA203497	Hs.124601	0	1	zx58g05.r1 cDNA, 5' end /clone=IMAGE:446744 /
517G2	577	756	AA858297	Hs.124675	3.00E-61	1	ob13b08.s1 cDNA, 3' end /clone=IMAGE:1323543
107H3	913	1220	AK023013	Hs.124762	1.00E-174	1	FLJ12951 fis, clone NT2RP2005457, highly
473A7	729	929	NM_019062	Hs.124835	4.00E-82	1	hypothetical protein (FLJ20225), mRNA /cds=(
108D12	3225	3531	AF023142	Hs.125134	1.00E-142	2	pre-mRNA splicing SR protein rA4 mRNA, partial
463E11	158	519	AI380443	Hs.125608	0	1	tg02f04.x1 cDNA, 3' end /clone=IMAGE:2107615
104F6	1581	2028	NM_019853	Hs.125682	0	1	protein phosphatase 4 regulatory subunit 2 (P
462A5	5	282	AW975851	Hs.125815	1.00E-149	1	EST387960 cDNA /gb=AW975851 /gi=8167072 /ug=
462B1	534	702	AI378032	Hs.125892	1.00E-69	1	te67g08.x1 cDNA, 3' end /clone=IMAGE:2091806
121A6	3074	3494	AB028978	Hs.126084	1.00E-174	1	mRNA for KIAA1055 protein, partial cds /cds=(0
171G12	94	1240	M15330	Hs.126256	0	7	interleukin 1-beta (IL1B) mRNA, complete cds /cds=(86
183D12	100	1275	NM_000576	Hs.126256	0	9	interleukin 1, beta (IL1B), mRNA /cds=(86,895)
458B2	6	415	AI393205	Hs.126265	0	1	tg14b07.x1 cDNA, 3' end /clone=IMAGE:2108725
102G6	885	1906	AJ271684	Hs.126355	1.00E-171	2	for myeloid DAP12-associating lectin (MD
463E4	847	1015	NM_013252	Hs.126355	2.00E-89	1	C-type (calcium dependent, carbohydrate-reco
167B2	2468	2721	AF195514	Hs.126550	1.00E-142	1	VPS4-2 ATPase (VPS42) mRNA, complete cds /cds=
473D8	19	397	BF445163	Hs.126594	0	1	nad21d12.x1 cDNA, 3' end /clone=IMAGE:3366191
143C9	333	551	BE250027	Hs.126701	1.00E-121	1	600943030F1 cDNA, 5' end /clone=IMAGE:2959639
471E10	806	945	AK021519	Hs.126707	2.00E-71	1	cDNA FLJ11457 fis, clone HEMBA1001522 /cds=(1
462B4	159	572	NM_017762	Hs.126721	0	1	hypothetical protein FLJ20313 (FLJ20313), mR
41D8	1	2519	AK023275	Hs.126925	0	5	FLJ13213 fis, clone NT2RP4001126, weakly
463F5	2	563	NM_014464	Hs.127011	0	1	tubulointerstitial nephritis antigen (TIN-A
597C8	2662	2905	AB046765	Hs.127270	1.00E-136	1	mRNA for KIAA1545 protein, partial cds /cds=(0
458F11	15	212	BF508731	Hs.127311	8.00E-81	1	UI-H-BI4-aoq-b-08-0-UI.s1 cDNA, 3' end /clon
462B3	76	389	AW978753	Hs.127327	1.00E-133	1	EST390862 cDNA /gb=AW978753 /gi=8170027 /ug=
463E2	176	787	AI028267	Hs.127514	0	1	ow01d06.x1 cDNA, 3' end /clone=IMAGE:1645547
465G5	181	372	AA953396	Hs.127557	6.00E-78	1	on63h10.s1 cDNA, 3' end /clone=IMAGE:1561411
463E10	11190	11634	NM_016239	Hs.127561	0	1	unconventional myosin-15 (LOC51168), mRNA /c
476A9	27	216	AW384918	Hs.127574	1.00E-101	1	PM1-HT0422-291299-002-d01 cDNA /gb=AW384918
111B10	1825	2463	NM_014007	Hs.127649	0	1	KIAA0414 protein (KIAA0414), mRNA /cds=(1132,
499A7	2134	5198	AF070674	Hs.127799	0	8	inhibitor of apoptosis protein-1 (MIHC) mRNA,
331F5	4	460	BF342439	Hs.127863	0	1	602013944F1 cDNA, 5' end /clone=IMAGE:4149562
176A12	796	1351	NM_022900	Hs.128003	0	1	hypothetical protein FLJ21213 (FLJ21213), mR
462B5	1766	1949	NM_014406	Hs.128342	5.00E-82	1	potassium large conductance calcium-activate

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

467D5	157	279	AI222805	Hs.128630	6.00E-62	1	qp39c07.x1 cDNA, 3' end /clone=IMAGE:1925388
465G3	1	529	BE222032	Hs.128675	0	1	hr61g11.x1 cDNA, 3' end /clone=IMAGE:3133028
467C7	1172	1726	AF118274	Hs.128740	0	1	DNb-5 mRNA, partial cds /cds=(0,1601) /gb=AF11
175G11	358	724	AL110151	Hs.128797	0	1	mRNA; cDNA DKFZp586D0824 (from clone DKFZp586
472A12	402	782	BE745645	Hs.129135	1.00E-153	1	601578727F1 cDNA, 5' end /clone=IMAGE:3927535
473C7	46	217	BE670584	Hs.129192	3.00E-37	1	7e36h08.x1 cDNA, 3' end /clone=IMAGE:3284607
463G11	7	397	AA746320	Hs.129572	0	1	ob08f01.s1 cDNA, 3' end /clone=IMAGE:1323097
63D8	18	1167	D13748	Hs.129673	0	4	eukaryotic initiation factor 4AI /cds=(16,12
57F3	19	1279	NM_001416	Hs.129673	0	4	eukaryotic translation initiation factor 4A,
144G5	1071	1192	AF064090	Hs.129708	3.00E-62	3	ligand for herpesvirus entry mediator (HVEM-L)
118A9	2684	3198	AB046805	Hs.129750	0	1	mRNA for KIAA1585 protein, partial cds /cds=(2
50G5	1119	1440	AK024068	Hs.129872	1.00E-172	1	FLJ14006 fis, clone Y79AA1002399, highly
469D6	376	603	D43968	Hs.129914	1.00E-126	1	AML1 mRNA for AML1b protein (alternatively spliced pr
590G11	823	1571	NM_003563	Hs.129951	0	3	speckle-type POZ protein (SPOP), mRNA /cds=(15
591C7	68	571	NM_005243	Hs.129953	0	1	Ewing sarcoma breakpoint region 1 (EWSR1), tra
459F5	579	768	AI763262	Hs.130059	1.00E-35	1	wi66c04.x1 cDNA, 3' end /clone=IMAGE:2398278
479A10	259	448	AI089359	Hs.130232	1.00E-103	1	qb05h03.x1 cDNA, 3' end /clone=IMAGE:1695413
461G5	193	347	AW898615	Hs.130729	2.00E-68	1	RC1-NN0073-090500-012-f02 cDNA /gb=AW898615
466B1	373	569	AI347054	Hs.130879	1.00E-76	1	qp60a04.x1 cDNA, 3' end /clone=IMAGE:1927374
463G3	3212	5430	AJ404611	Hs.130881	0	2	mRNA for B-cell lymphoma/leukaemia 11A extra
462C3	48	468	AI421806	Hs.131067	0	1	tf44h11.x1 cDNA, 3' end /clone=IMAGE:2099109
596G10	39	491	NM_006294	Hs.131255	0	3	ubiquinol-cytochrome c reductase binding pro
469G10	189	361	AI024984	Hs.131580	1.00E-81	1	ov39d11.x1 cDNA, 3' end /clone=IMAGE:1639701
458B7	169	659	AW978870	Hs.131828	0	1	EST390979 cDNA /gb=AW978870 /gi=8170147 /ug=
63D1	185	500	AF176706	Hs.131859	1.00E-133	1	F-box protein FBX11 mRNA, partial cds /cds=(0,
58C10	4188	4313	NM_014913	Hs.131915	2.00E-65	1	KIAA0863 protein (KIAA0863), mRNA /cds=(185,3
117H2	282	569	NM_003608	Hs.131924	1.00E-143	1	G protein-coupled receptor 65 (GPR65), mRNA /
462D11	441	683	AW976422	Hs.132064	1.00E-118	1	EST388531 cDNA /gb=AW976422 /gi=8167649 /ug=
586F11	161	1094	NM_017830	Hs.132071	0	2	hypothetical protein FLJ20455 (FLJ20455), mR
466A8	118	224	AI042377	Hs.132156	2.00E-44	1	ox62c03.x1 cDNA, 3' end /clone=IMAGE:1660900
472F6	979	1431	AK022463	Hs.132221	0	1	cDNA FLJ12401 fis, clone MAMMA1002796 /cds=(3,
462E4	19	567	AI031656	Hs.132237	0	1	ow48e06.x1 cDNA, 3' end /clone=IMAGE:1650082
462E2	4	539	AI829569	Hs.132238	0	1	wf28e02.x1 cDNA, 3' end /clone=IMAGE:2356922
461H9	453	618	BG037042	Hs.132555	4.00E-57	1	602288311F1 cDNA, 5' end /clone=IMAGE:4374122
467D10	4518	4689	AK024449	Hs.132569	2.00E-55	1	mRNA for FLJ00041 protein, partial cds /cds=(0
463H7	162	438	AI346336	Hs.132594	1.00E-132	1	qp50b04.x1 cDNA, 3' end /clone=IMAGE:1926415
592B8	2415	2957	NM_005337	Hs.132834	0	1	hematopoietic protein 1 (HEM1), mRNA /cds=(158
70H2	6370	6718	AF047033	Hs.132904	1.00E-175	1	sodium bicarbonate cotransporter 3 (SLC4A7) m
50G10	1167	2041	AL121985	Hs.132906	0	4	DNA sequence from clone RP11-404F10 on chromosome 1q2
123C10	1323	1570	NM_015071	Hs.132942	1.00E-136	1	GTPase regulator associated with the focal adh
121B10	92	503	AA504269	Hs.133032	0	1	aa61c09.s1 cDNA, 3' end /clone=IMAGE:825424 /
171A12	696	909	AL050035	Hs.133130	6.00E-83	1	mRNA; cDNA DKFZp566H0124 (from clone DKFZp566
463B5	123	449	AI051673	Hs.133175	1.00E-176	1	oy77g06.x1 cDNA, 3' end /clone=IMAGE:1671898
463B7	966	1103	AL044498	Hs.133262	3.00E-46	1	DKFZp434I082_s1 cDNA, 3' end /clone=DKFZp434I
463B8	1	322	AV661783	Hs.133333	1.00E-176	1	AV661783 cDNA, 3' end /clone=GLCGXE12 /clone_
463A10	431	694	AW966876	Hs.133543	1.00E-110	1	EST378950 cDNA /gb=AW966876 /gi=8156712 /ug=
464B10	63	547	BF965766	Hs.133864	0	1	602276890F1 cDNA, 5' end /clone=IMAGE:4364495

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

460C6	454	653	AW009671	Hs.134272	8.00E-70	1	ws85g09.x1 cDNA, 3' end /clone=IMAGE:2504800
459C12	3337	3745	AJ278245	Hs.134342	1.00E-121	1	mRNA for LanC-like protein 2 (lancl2 gene) /cds
462G1	33	454	AI074016	Hs.134473	0	1	oy66g02.x1 cDNA, 3' end /clone=IMAGE:1670834
462G6	260	597	BE676210	Hs.134648	1.00E-156	1	7f25c05.x1 cDNA, 3' end /clone=IMAGE:3295688
466H12	505	662	AV706481	Hs.134829	3.00E-65	1	AV706481 cDNA, 5' end /clone=ADBBYF02
148H11	16	474	BE786820	Hs.135056	0	1	601477630F1 5' end /clone=IMAGE:3880471
462E1	139	487	BF109873	Hs.135106	0	1	7l70e11.x1 cDNA, 3' end /clone=IMAGE:3526772
147E6	11	364	AV712376	Hs.135167	0	2	AV712376 cDNA, 5' end /clone=DCAAND12 /clone_
465B4	1993	2237	AJ271326	Hs.135187	1.00E-92	1	mRNA for unc-93 related protein (UNC93 gene) /
463B4	185	352	AI051664	Hs.135339	4.00E-48	1	oy77f06.x1 cDNA, 3' end /clone=IMAGE:1671875
478H4	2126	2458	AK024921	Hs.135570	1.00E-170	1	cDNA: FLJ21268 fis, clone COL01718 /cds=UNKNOW
148B6	119	444	AI004582	Hs.135764	3.00E-82	8	ou04a11.x1 3' end /clone=IMAGE:1625276
598E9	1948	2184	NM_022117	Hs.136164	3.00E-93	1	cutaneous T-cell lymphoma-associated tumor a
514C10	398	840	AL049597	Hs.136309	0	2	DNA sequence from clone RP4-612B15 on chromosome 1p22
461C6	18	219	BF513274	Hs.136375	1.00E-101	1	UI-H-BW1-amo-d-11-0-UI.s1 cDNA, 3' end /clon
482E4	291	699	BF526066	Hs.136537	1.00E-142	1	602071176F1 cDNA, 5' end /clone=IMAGE:4214059
461G7	43	466	NM_013378	Hs.136713	0	1	pre-B lymphocyte gene 3 (VPREB3), mRNA /cds=(4
119B10	10	677	NM_013269	Hs.136748	0	2	lectin-like NK cell receptor (LLT1), mRNA /cd
462A10	1233	1727	AK024426	Hs.137354	0	1	mRNA for FLJ00015 protein, partial cds /cds=(3
41F2	2684	3000	AJ223324	Hs.137548	1.00E-156	1	for MAX.3 cell surface antigen /cds=(44,10
74E8	16	2000	D10923	Hs.137555	0	15	HM74 /cds=(60,1223) /gb=D10923 /gi=219866 /
58D10	8	2000	NM_006018	Hs.137555	0	9	putative chemokine receptor; GTP-binding pro
120E2	210	814	NM_002027	Hs.138381	0	1	farnesyltransferase, CAAX box, alpha (FNTA),
168E12	1953	2522	D38524	Hs.138593	0	1	5'-nucleotidase /cds=(83,1768) /gb=D38524
178F7	573	824	NM_006413	Hs.139120	1.00E-115	1	ribonuclease P (30kD) (RPP30), mRNA /cds=(27,8
473D1	1635	1767	AL049942	Hs.139240	6.00E-50	1	mRNA; cDNA DKFZp564F1422 (from clone DKFZp564F
188A8	924	1038	NM_017523	Hs.139262	1.00E-56	2	XIAP associated factor-1 (HSXIAPAF1), mRNA /c
168F7	933	1038	X99699	Hs.139262	1.00E-53	1	for XIAP associated factor-1 /cds=(0,953) /
181B10	1556	2517	NM_005816	Hs.142023	0	3	T cell activation, increased late expression (
514E7	2052	2339	NM_003150	Hs.142258	1.00E-114	1	signal transducer and activator of transcripti
196C7	355	524	NM_016123	Hs.142295	9.00E-92	1	putative protein kinase NY-REN-64 antigen (LO
585B10	3261	3465	AK023129	Hs.142442	1.00E-100	1	cDNA FLJ13067 fis, clone NT2RP3001712, highly
458F2	283	413	BE293343	Hs.142737	3.00E-68	1	601143756F1 cDNA, 5' end /clone=IMAGE:3051493
134C6	289	572	BE886127	Hs.142838	1.00E-160	1	601509912F1 cDNA, 5' end /clone=IMAGE:3911451
110A11	345	584	AI126688	Hs.143049	1.00E-102	1	qb94a06.x1 cDNA, 3' end /clone=IMAGE:1707730
472G7	127	452	AW976331	Hs.143254	0	1	EST388440 cDNA /gb=AW976331 /gi=8167557 /ug=
464G11	425	547	AI357640	Hs.143314	1.00E-56	1	qy15b06.x1 cDNA, 3' end /clone=IMAGE:2012051
463F11	257	640	BF446017	Hs.143389	0	1	7p18a11.x1 cDNA, 3' end /clone=IMAGE:3646004
463H2	107	443	AA825245	Hs.143410	1.00E-151	1	oe59g09.s1 cDNA, 3' end /clone=IMAGE:1415968
48B7	1	3366	NM_005813	Hs.143460	0	2	protein kinase C, nu (PRKCN), mRNA /cds=(555,32
463C9	290	405	AW173163	Hs.143525	5.00E-41	1	xj84b08.x1 cDNA, 3' end /clone=IMAGE:2663895
463C8	330	473	AI095189	Hs.143534	5.00E-57	2	oy83b06.s1 cDNA, 3' end /clone=IMAGE:1672403
464G5	94	189	BG033028	Hs.143554	1.00E-38	1	602300135F1 cDNA, 5' end /clone=IMAGE:4401776
463D7	120	563	NM_006777	Hs.143604	0	1	Kaiso (ZNF-kaiso), mRNA /cds=(0,2018) /gb=NM
471A10	132	586	AK026372	Hs.143631	0	1	cDNA: FLJ22719 fis, clone HSI14307 /cds=UNKNOW
74G2	5129	5285	AF073310	Hs.143648	2.00E-79	2	insulin receptor substrate-2 (IRS2) mRNA, com
471G11	7	320	AI568622	Hs.143951	1.00E-154	2	tn41e10.x1 cDNA, 3' end /clone=IMAGE:2170218
478H12	963	1532	NM_018270	Hs.143954	0	1	hypothetical protein FLJ10914 (FLJ10914), mR
462G3	100	529	AI074020	Hs.144114	0	1	oy66g06.x1 cDNA, 3' end /clone=IMAGE:1670842
463C1	52	151	AI090305	Hs.144119	1.00E-42	1	oy81b01.s1 cDNA, 3' end /clone=IMAGE:1672201

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

472H8	157	485	BF509758	Hs.144265	1.00E-178	1	UI-H-BI4-apg-d-04-0-UI.s1 cDNA, 3' end /clon
166E1	23	443	D63874	Hs.144321	0	1	HMG-1, complete cds /cds=(76,723) /gb=D63874
145G8	125	1606	NM_018548	Hs.144477	0	2	hypothetical protein PRO2975 (PRO2975), mRNA
191H8	46	624	BF036686	Hs.144559	0	1	601459771F1 cDNA, 5' end /clone=IMAGE:3863248
151B1	1983	2561	M93651	Hs.145279	0	2	set gene, complete cds /cds=(3,836) /gb=M93651 /gi=33
514B2	115	1583	NM_003011	Hs.145279	0	4	SET translocation (myeloid leukemia-associat
596D4	89	734	AA631938	Hs.145668	0	8	fmfc5 cDNA /clone=CR6-21 /gb=AA631938 /gi=25
492B3	512	2226	NM_004902	Hs.145696	0	2	splicing factor (CC1.3) (CC1.3), mRNA /cds=(14
192E4	1483	1837	AF246126	Hs.145956	0	1	zinc finger protein mRNA, complete cds /cds=(1
480B9	1094	1426	AL136874	Hs.146037	1.00E-111	1	mRNA; cDNA DKFZp434C135 (from clone DKFZp434C1
49H1	1761	2182	NM_022894	Hs.146123	0	1	hypothetical protein FLJ12972 (FLJ12972), mR
129C6	517	603	BE220959	Hs.146215	6.00E-21	1	hu02b06.x1 cDNA, 3' end /clone=IMAGE:3165395
583D9	249	646	NM_003641	Hs.146360	0	1	interferon induced transmembrane protein 1 (
589D9	125	1866	NM_002139	Hs.146381	0	5	RNA binding motif protein, X chromosome (RBMX)
68H11	122	1567	Z23064	Hs.146381	0	2	mRNA gene for hnRNP G protein /cds=(11,1186) /gb=
174A8	461	1008	NM_004757	Hs.146401	0	1	small inducible cytokine subfamily E, member 1
171A6	461	686	U10117	Hs.146401	1.00E-100	1	endothelial-monocyte activating polypeptide II mRN
465C4	53	342	AI141004	Hs.146627	3.00E-89	1	oy68f02.x1 cDNA, 3' end /clone=IMAGE:1671003
190H7	1306	3107	AB033079	Hs.146668	0	3	mRNA for KIAA1253 protein, partial cds /cds=(0
102E9	412	1022	AF054187	Hs.146763	0	3	alpha NAC mRNA, complete cds /cds=(309,956) /g
179B1	364	843	D16481	Hs.146812	0	1	mitochondrial 3-ketoacyl-CoA thiolas
126H12	1	358	NM_000183	Hs.146812	0	1	hydroxyacyl-Coenzyme A dehydrogenase/3-keto
476C9	20	249	AI187423	Hs.147040	1.00E-128	2	qf31d04.x1 cDNA, 3' end /clone=IMAGE:1751623
70H11	47	1593	AF272148	Hs.147644	0	7	KRAB zinc finger protein (RITA) mRNA, complete
51F1	635	1039	NM_018555	Hs.147644	0	3	C2H2-like zinc finger protein (ZNF361), mRNA
72H1	948	5026	AF000982	Hs.147916	0	7	dead box, X isoform (DBX) mRNA, alternative tra
37F10	3128	3652	X63563	Hs.148027	0	1	RNA polymerase II 140 kDa /cds=(43,3567)
64C11	163	279	AA908367	Hs.148288	6.00E-29	1	og76c11.s1 cDNA, 3' end /clone=IMAGE:1454228
463G2	52	473	AI335004	Hs.148558	0	1	tb21e09.x1 cDNA, 3' end /clone=IMAGE:2055016
471F8	17	463	AI471866	Hs.149095	0	1	ti67d04.x1 cDNA, 3' end /clone=IMAGE:2137063
169C12	449	1711	L06132	Hs.149155	0	2	voltage-dependent anion channel isoform 1 (VDAC) mRN
189G6	1353	1711	NM_003374	Hs.149155	0	5	voltage-dependent anion channel 1 (VDAC1), mR
481E3	501	669	NM_007022	Hs.149443	5.00E-84	1	putative tumor suppressor (101F6), mRNA /cds=
472B3	93	182	BF029894	Hs.149595	6.00E-44	1	601557056F1 cDNA, 5' end /clone=IMAGE:3827172
173D1	3719	3877	AB037901	Hs.149918	3.00E-83	1	GASC-1 mRNA, complete cds /cds=(150,3320) /gb
153G12	1429	1787	M31627	Hs.149923	0	2	X box binding protein-1 (XBP-1) mRNA, complete cds /cd
116B10	1435	1787	NM_005080	Hs.149923	1.00E-180	1	X-box binding protein 1 (XBP1), mRNA /cds=(12,7
111G4	480	1891	L12052	Hs.150395	0	2	cAMP phosphodiesterase PDE7 (PDE7A1) mRNA, co
461D6	1407	1904	NM_000790	Hs.150403	0	1	dopa decarboxylase (aromatic L-amino acid dec
73B3	896	1779	AL050005	Hs.150580	0	23	cDNA DKFZp564A153 (from clone DKFZp564A1
465G12	1	549	AJ272212	Hs.150601	0	1	mRNA for protein serine kinase (PSKH1 gene) /c
140G12	2	195	BF028489	Hs.150675	1.00E-100	1	601763692F1 cDNA, 5' end /clone=IMAGE:3995950
496E10	17	1686	BC000167	Hs.151001	0	5	clone IMAGE:2900671, mRNA, partial cds /cds=
597G7	623	1488	NM_005015	Hs.151134	0	2	oxidase (cytochrome c) assembly 1-like (OXA1L
50C9	1051	1467	X80695	Hs.151134	0	1	OXA1Hs mRNA /cds=(6,1313) /gb=X80695 /gi=619490
125H7	3154	3957	NM_001421	Hs.151139	0	3	E74-like factor 4 (ets domain transcription fa
111F2	306	638	BG286500	Hs.151239	1.00E-149	1	602382992F1 cDNA, 5' end /clone=IMAGE:4500527
177A4	9686	10035	AF075587	Hs.151411	0	1	protein associated with Myc mRNA, complete cds
185C7	6934	13968	NM_015057	Hs.151411	0	3	KIAA0916 protein (KIAA0916), mRNA /cds=(146,1
115E7	3406	4005	NM_004124	Hs.151413	0	1	glia maturation factor, beta (GMFB), mRNA /cds

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

182H7	234	833	AF099032	Hs.151461	0	1	embryonic ectoderm development protein short
169C10	4247	4727	U38847	Hs.151518	0	1	TAR RNA loop binding protein (TRP-185) mRNA, complete
167D6	1013	1197	NM_002870	Hs.151536	6.00E-83	1	RAB13, member RAS oncogene family (RAB13), mRN
588G11	1249	1898	AK023362	Hs.151604	1.00E-157	9	cDNA FLJ13300 fis, clone OVARC1001342, highly
479G10	1	277	NM_007210	Hs.151678	1.00E-103	1	UDP-N-acetyl-alpha-D-galactosamine:polype
178B7	2664	3033	NM_004247	Hs.151787	0	4	U5 snRNP-specific protein, 116 kD (U5-116KD),
59A6	382	860	D42054	Hs.151791	0	1	KIAA0092 gene, complete cds /cds=(53,1477) /
521B6	2017	2205	NM_014679	Hs.151791	2.00E-93	1	KIAA0092 gene product (KIAA0092), mRNA /cds=(
59C10	37	697	AF070525	Hs.151903	0	5	clone 24706 mRNA sequence /cds=UNKNOWN /gb=AF
519A7	165	686	NM_005792	Hs.152720	0	1	M-phase phosphoprotein 6 (MPHOSPH6), mRNA /c
481E11	3990	4280	NM_005154	Hs.152818	1.00E-135	1	ubiquitin specific protease 8 (USP8), mRNA /cd
110F2	1210	1841	L25931	Hs.152931	0	2	lamin B receptor (LBR) mRNA, complete cds /cds=(75,192
516F8	1217	1708	NM_002296	Hs.152931	0	1	lamin B receptor (LBR), mRNA /cds=(75,1922) /g
462B2	93	2385	AF244129	Hs.153042	0	2	cell-surface molecule Ly-9 mRNA, complete cds
41F4	617	905	X14046	Hs.153053	1.00E-162	1	leukocyte antigen CD37 /cds=(63,908) /gb=X14
462G8	2312	2843	AF311312	Hs.153057	0	1	infertility-related sperm protein mRNA, comp
142H5	17	221	M94856	Hs.153179	1.00E-92	1	fatty acid binding protein homologue (PA-FABP) mRNA,
486G9	3	431	NM_001444	Hs.153179	0	1	fatty acid binding protein 5 (psoriasis-associ
40A1	2158	2716	X79201	Hs.153221	0	1	SYT /cds=(3,1178) /gb=X79201 /gi=531105
101D9	1524	2060	AB014601	Hs.153293	0	1	for KIAA0701 protein, partial cds /cds=(0
460F10	1457	6107	AB032972	Hs.153489	0	2	mRNA for KIAA1146 protein, partial cds /cds=(0
106A5	445	547	AI761622	Hs.153523	2.00E-37	1	wg66f05.x1 cDNA, 3' end /clone=IMAGE:2370081
482A6	49	369	AI859076	Hs.153551	1.00E-106	1	wl33b04.x1 cDNA, 3' end /clone=IMAGE:2426671
589B2	1054	1556	AF261091	Hs.153612	0	1	iron inhibited ABC transporter 2 mRNA, complet
57A3	1586	1757	NM_004073	Hs.153640	9.00E-87	1	cytokine-inducible kinase (CNK), mRNA /cds=(3
466H3	2	257	NM_003866	Hs.153687	1.00E-133	1	inositol polyphosphate-4-phosphatase, type
483B6	3337	3544	NM_002526	Hs.153952	2.00E-72	1	5' nucleotidase (CD73) (NT5), mRNA /cds=(49,17
41F1	2749	3371	X55740	Hs.153952	0	1	placental cDNA coding for 5'nucleotidase (EC 3.1.3.5)
44C3	1319	1574	X82206	Hs.153961	1.00E-130	1	alpha-centractin /cds=(66,1196) /gb=X8
64F12	2578	2713	NM_022790	Hs.154057	1.00E-26	1	matrix metalloproteinase 19 (MMP19), transcri
72E11	1886	2717	U38320	Hs.154057	0	15	clone rasi-3 matrix metalloproteinase RASI-1
165H12	414	663	AW970676	Hs.154172	2.00E-22	1	EST382759 cDNA /gb=AW970676 /gi=8160521 /ug=
37A4	1151	2746	M31210	Hs.154210	0	2	endothelial differentiation protein (edg-1) gene mR
597F4	1125	2395	NM_001400	Hs.154210	0	11	endothelial differentiation, sphingolipid G
106F2	24	1657	U22897	Hs.154230	0	2	nuclear domain 10 protein (ndp52) mRNA, comple
466E2	116	373	AB023149	Hs.154296	1.00E-131	2	mRNA for KIAA0932 protein, partial cds /cds=(0
107F11	1386	1743	AL117566	Hs.154320	0	1	cDNA DKFZp566J164 (from clone DKFZp566J1
166E12	4490	4894	D86967	Hs.154332	0	1	KIAA0212 gene, complete cds /cds=(58,2031) /
188D12	5148	5666	NM_014674	Hs.154332	0	2	KIAA0212 gene product (KIAA0212), mRNA /cds=(
66A1	88	615	M82882	Hs.154365	0	1	cis-acting sequence /cds=UNKNOWN /gb=M82882 /gi=180
37C1	4320	4776	AB028999	Hs.154525	0	1	for KIAA1076 protein, partial cds /cds=(0
98D2	2317	4907	NM_000104	Hs.154654	0	6	cytochrome P450, subfamily I (dioxin-inducibl
37C4	4445	4907	U03688	Hs.154654	0	3	dioxin-inducible cytochrome P450 (CYP1B1) mRNA, comp
464A5	1418	2027	NM_006636	Hs.154672	0	3	methylene tetrahydrofolate dehydrogenase (N
36C5	615	1689	X16396	Hs.154672	0	7	NAD-dependent methylene tetrahydrofolate d
67C8	1	397	U85773	Hs.154695	0	1	phosphomannomutase (PMM2) mRNA, complete cds /cds=(
525D3	2084	2533	NM_002651	Hs.154846	0	1	phosphatidylinositol 4-kinase, catalytic, b
109A7	1979	3148	D10040	Hs.154890	0	2	for long-chain acyl-CoA synthetase, compl
167F6	1817	3359	NM_021122	Hs.154890	0	8	fatty-acid-Coenzyme A ligase, long-chain 2 (

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

182A1	344	793	NM_021825	Hs.154938	0	1	hypothetical protein MDS025 (MDS025), mRNA /
104E2	1254	1762	D87450	Hs.154978	0	1	KIAA0261 gene, partial cds /cds=(0,3865) /gb
519G10	4912	5303	NM_003489	Hs.155017	0	1	nuclear receptor interacting protein 1 (NRIP1
595C6	4067	4631	NM_006526	Hs.155040	0	2	zinc finger protein 217 (ZNF217), mRNA /cds=(2
105D4	1768	2418	L42373	Hs.155079	0	1	phosphatase 2A B56-alpha (PP2A) mRNA, complete
174B7	1768	2320	NM_006243	Hs.155079	0	1	protein phosphatase 2, regulatory subunit B (
75G4	920	1775	X59066	Hs.155101	0	2	mitochondrial ATP synthase (F1-ATPase) alpha
523G12	20	848	NM_004681	Hs.155103	0	3	eukaryotic translation initiation factor 1A,
74D7	292	1094	M16942	Hs.155122	0	3	MHC class II HLA-DRw53-associated glycoprotein beta-
137D4	2500	2822	AL049761	Hs.155140	1.00E-176	1	DNA sequence from clone RP5-863C7 on chromosome 20p12
471B5	908	1168	AK023379	Hs.155160	1.00E-141	1	cDNA FLJ13317 fis, clone OVARC1001577, highly
176C9	2104	2635	NM_003664	Hs.155172	0	1	adaptor-related protein complex 3, beta 1 sub
99F5	212	671	NM_005642	Hs.155188	0	1	TATA box binding protein (TBP)-associated fac
166E9	1215	1637	U18062	Hs.155188	0	1	TFIID subunit TAFII55 (TAFII55) mRNA, complete cds /c
163A11	60	3052	AL162086	Hs.155191	0	8	cDNA DKFZp762H157 (from clone DKFZp762H1
71E4	44	558	NM_003379	Hs.155191	1.00E-175	4	villin 2 (ezrin) (VIL2), mRNA /cds=(117,1877)
145D8	2135	2669	L47345	Hs.155202	0	1	elongin A mRNA, complete cds /cds=(32,2350) /g
477H9	357	2812	NM_014670	Hs.155291	0	2	KIAA0005 gene product (KIAA0005), mRNA /cds=(
58D8	38	336	NM_000518	Hs.155376	1.00E-100	1	hemoglobin, beta (HBB), mRNA /cds=(50,493) /g
48F11	576	2131	NM_006164	Hs.155396	0	2	nuclear factor (erythroid-derived 2)-like 2
65G11	426	1179	S74017	Hs.155396	0	1	Nrf2=NF-E2-like basic leucine zipper transcriptional act
480G12	852	1246	NM_001352	Hs.155402	0	1	D site of albumin promoter (albumin D-box) bind
182B12	245	592	NM_006899	Hs.155410	0	1	isocitrate dehydrogenase 3 (NAD+) beta (IDH3B
599C9	3188	3487	NM_021643	Hs.155418	1.00E-163	1	GS3955 protein (GS3955), mRNA /cds=(1225,2256
68H2	563	1749	AF037448	Hs.155489	0	2	RRM RNA binding protein Gry-rbp (GRY-RBP) mRNA
173F6	1243	1811	AF208043	Hs.155530	0	2	IFI16b (IFI16b) mRNA, complete cds /cds=(264,2
170B3	1061	1342	D50063	Hs.155543	1.00E-139	1	proteasome subunit p40_/ Mov34 protein, comp
590E9	494	1323	NM_002811	Hs.155543	0	2	proteasome (prosome, macropain) 26S subunit,
522D11	1463	1710	AB029003	Hs.155546	1.00E-138	2	mRNA for KIAA1080 protein, partial cds /cds=(0
587A8	3514	3923	NM_001746	Hs.155560	0	1	calnexin (CANX), mRNA /cds=(89,1867) /gb=NM_0
39A6	830	1474	D63878	Hs.155595	0	1	KIAA0158 gene, complete cds /cds=(258,1343)
167F5	745	2735	NM_004404	Hs.155595	0	3	neural precursor cell expressed, developmenta
106E10	1922	2340	U15173	Hs.155596	1.00E-179	2	BCL2/adenovirus E1B 19kD-interacting protein
524A8	1639	2229	NM_014666	Hs.155623	0	1	KIAA0171 gene product (KIAA0171), mRNA /cds=(
166D6	12177	12974	U47077	Hs.155637	0	3	DNA-dependent protein kinase catalytic subuni
488A10	1961	2426	NM_002827	Hs.155894	0	3	protein tyrosine phosphatase, non-receptor t
65D6	696	1107	S68271	Hs.155924	0	3	cyclic AMP-responsive element modulator (CRE
113E8	682	1435	NM_004054	Hs.155935	0	1	complement component 3a receptor 1 (C3AR1), mR
105F10	119	1591	U62027	Hs.155935	0	3	anaphylatoxin C3a receptor (HNFAG09) mRNA, complete
111C1	4122	4779	NM_005541	Hs.155939	0	5	inositol polyphosphate-5-phosphatase, 145kD
40A9	1727	2300	D76444	Hs.155968	0	1	hkf-1 mRNA, complete cds /cds=(922,2979) /gb=
124F1	1464	2121	NM_005667	Hs.155968	0	1	zinc finger protein homologous to Zfp103 in mo
481E12	2237	2691	NM_003588	Hs.155976	0	1	cullin 4B (CUL4B), mRNA /cds=(78,2231) /gb=NM
109H3	36	440	NM_020414	Hs.155986	0	1	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide
193B10	1103	1892	AK024974	Hs.156110	1.00E-180	5	cDNA: FLJ21321 fis, clone COL02335, highly sim
463H6	26	149	AI337347	Hs.156339	5.00E-57	1	tb98e10.x1 cDNA, 3' end /clone=IMAGE:2062410
107H5	34	253	AI146787	Hs.156601	7.00E-93	1	qb83f02.x1 cDNA, 3' end /clone=IMAGE:1706715
517E8	209	822	NM_015646	Hs.156764	0	3	RAP1B, member of RAS oncogene family (RAP1B),
478H11	456	768	NM_005819	Hs.157144	1.00E-172	1	syntaxin 6 (STX6), mRNA /cds=(0,767) /gb=NM_0

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

463G12	44	283	AI351144	Hs.157213	3.00E-95	1	qt23f10.x1 cDNA, 3' end /clone=IMAGE:1948459
520A2	2359	2565	BC001913	Hs.157236	1.00E-95	2	Similar to membrane protein of cholinergic sy
473A8	2944	3570	AK026394	Hs.157240	0	1	cDNA: FLJ22741 fis, clone HUV00774 /cds=UNKNOW
464D5	433	601	AW207701	Hs.157315	8.00E-37	1	UI-H-BI2-age-e-03-0-UI.s1 cDNA, 3' end /clon
464B8	288	633	BF184881	Hs.157396	2.00E-99	1	601843756F1 cDNA, 5' end /clone=IMAGE:4064508
463A6	225	554	AW976630	Hs.157447	1.00E-169	1	EST388739 cDNA /gb=AW976630 /gi=8167861 /ug=
464G10	423	661	AI356405	Hs.157556	1.00E-103	1	qz26g04.x1 cDNA, 3' end /clone=IMAGE:2028054
464H3	396	642	AI568755	Hs.157564	1.00E-123	1	th15f03.x1 cDNA, 3' end /clone=IMAGE:2118365
466C1	110	384	AI760026	Hs.157569	1.00E-135	1	wh83c05.x1 cDNA, 3' end /clone=IMAGE:2387336
465A2	11	178	AI823541	Hs.157710	1.00E-79	1	wh55c11.x1 cDNA, 3' end /clone=IMAGE:2384660
464A8	2000	2248	AK023779	Hs.157777	1.00E-134	1	cDNA FLJ13717 fis, clone PLACE2000425 /cds=UNK
464G1	122	447	AI361761	Hs.157813	1.00E-163	2	qz19a07.x1 cDNA, 3' end /clone=IMAGE:2021940
464G7	293	395	AI361849	Hs.157815	4.00E-30	1	qz19h11.x1 cDNA, 3' end /clone=IMAGE:2022021
145B8	238	598	BF303931	Hs.157850	1.00E-179	3	601886564F2 cDNA, 5' end /clone=IMAGE:4120574
115D1	111	712	NM_000661	Hs.157850	1.00E-159	2	ribosomal protein L9 (RPL9), mRNA /cds=(29,607
102F8	4161	4818	AB023198	Hs.158135	0	1	for KIAA0981 protein, partial cds /cds=(0
597H12	1253	2625	NM_000593	Hs.158164	0	5	ATP-binding cassette, sub-family B (MDR/TAP),
465A3	172	342	T78173	Hs.158193	5.00E-64	1	yd79c05.r1 cDNA, 5' end /clone=IMAGE:114440 /
465H8	740	1171	NM_006354	Hs.158196	1.00E-149	1	transcriptional adaptor 3 (ADA3, yeast homolo
59H12	1646	6883	NM_002313	Hs.158203	0	4	actin-binding LIM protein (ABLIM), transcript
464A2	32	549	NM_004571	Hs.158225	0	1	PBX/knotted 1 hoemobox 1 (PKNOX1), mRNA /cds=(
124F12	6603	6907	AB007915	Hs.158286	1.00E-172	1	mRNA for KIAA0446 protein, partial cds /cds=(3
519F5	80	268	AI199223	Hs.158289	1.00E-86	1	qi47c06.x1 cDNA, 3' end /clone=IMAGE:1859626
463F8	33	286	BF433857	Hs.158501	1.00E-123	1	7q71b07.x1 cDNA /clone=IMAGE /gb=BF433857 /g
137A8	204	452	AI370965	Hs.158653	5.00E-32	1	ta29b11.x1 cDNA, 3' end /clone=IMAGE:2045469
466A11	1	565	BE676408	Hs.158714	0	1	7f29b11.x1 cDNA, 3' end /clone=IMAGE:3296061
73C2	5	396	AW362008	Hs.158794	0	1	PM2-CT0265-211099-002-d04 /gb=AW362008
465C6	242	433	AI378113	Hs.158877	2.00E-95	1	tc80c12.x1 cDNA, 3' end /clone=IMAGE:2072470
465C2	29	153	AI378457	Hs.158894	4.00E-60	2	tc79d10.x1 cDNA, 3' end /clone=IMAGE:2072371
465C10	47	442	AI379953	Hs.158943	0	1	tc81a07.x1 cDNA, 3' end /clone=IMAGE:2072532
477B9	151	396	AI380220	Hs.158965	1.00E-109	2	tf94a04.x1 cDNA, 3' end /clone=IMAGE:2106894
477B10	1	414	AI380236	Hs.158966	0	2	tf94b10.x1 cDNA, 3' end /clone=IMAGE:2106907
466F8	128	233	AI380388	Hs.158975	4.00E-30	1	tf96a03.x1 cDNA, 3' end /clone=IMAGE:2107084
467E12	109	350	AI799909	Hs.158989	1.00E-82	1	wc46c08.x1 cDNA, 3' end /clone=IMAGE:2321678
469G6	169	470	AI631850	Hs.158992	1.00E-119	1	wa36h07.x1 cDNA, 3' end /clone=IMAGE:2300221
467H4	17	292	BF508694	Hs.158999	1.00E-117	1	UI-H-BI4-aop-f-09-0-UI.s1 cDNA, 3' end /clon
469B2	179	388	AI568751	Hs.159014	4.00E-94	1	th15d09.x1 cDNA, 3' end /clone=IMAGE:2118353
464E8	742	945	AL538276	Hs.159065	1.00E-110	1	AL538276 cDNA /clone=CS0DF027YC09-(5-prime)
469D9	1	413	AI431873	Hs.159103	0	1	ti26b11.x1 cDNA, 3' end /clone=IMAGE:2131581
122C7	1916	2375	NM_003266	Hs.159239	0	1	toll-like receptor 4 (TLR4), mRNA /cds=(284,26
462H4	79	239	BF307871	Hs.159336	7.00E-66	1	601890687F1 cDNA, 5' end /clone=IMAGE:4132028
179C1	428	734	AJ225093	Hs.159386	3.00E-88	1	single-chain antibody, complete cds
473D11	267	339	AI380255	Hs.159424	5.00E-34	1	tf94d08.x1 cDNA, 3' end /clone=IMAGE:2106927
107B2	1	617	BE783628	Hs.159441	1.00E-160	2	601471696F1 cDNA, 5' end /clone=IMAGE:3874823
590E12	52	654	BG290141	Hs.159441	0	6	602385221F1 cDNA, 5' end /clone=IMAGE:4514380
70E1	2095	2333	AK027194	Hs.159483	1.00E-119	1	FLJ23541 fis, clone LNG08276, highly sim
58A5	10448	12675	AF193556	Hs.159492	0	10	sacsin (SACS) gene, complete cds /cds=(76,1156
482E11	2064	2559	NM_000061	Hs.159494	0	1	Bruton agammaglobulinemia tyrosine kinase (B
147A11	755	2415	AF001622	Hs.159523	0	7	class-I MHC-restricted T cell associated mole
486H6	1164	1382	NM_019604	Hs.159523	1.00E-117	2	class-I MHC-restricted T cell associated mole
465A5	2693	3039	NM_000033	Hs.159546	1.00E-148	1	ATP-binding cassette, sub-family D (ALD), mem

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

60C4	1102	1962	AK024833	Hs.159557	1.00E-147	4	FLJ21180 fis, clone CAS11176, highly sim
465B11	457	1126	NM_016952	Hs.159565	0	1	surface glycoprotein, Ig superfamily member (
477A12	89	581	AI797788	Hs.159577	0	5	wh78b11.x1 cDNA, 3' end /clone=IMAGE:2386845
595H8	19	912	NM_004632	Hs.159627	0	2	death associated protein 3 (DAP3), mRNA /cds=(
74D2	7	2119	AF153609	Hs.159640	0	9	serine/threonine protein kinase sgk mRNA, com
71B2	8	533	NM_005627	Hs.159640	0	1	serum/glucocorticoid regulated kinase (SGK)
467G8	310	488	AW006352	Hs.159643	2.00E-92	1	wt04d12.x1 cDNA, 3' end /clone=IMAGE:2506487
467B8	11	363	AI392893	Hs.159655	1.00E-173	1	tg05d07.x1 cDNA, 3' end /clone=IMAGE:2107885
471F11	16	303	AI827950	Hs.159659	1.00E-162	1	wk31a11.x1 cDNA, 3' end /clone=IMAGE:2413916
467C11	18	501	BF508053	Hs.159673	0	1	UI-H-BI4-apx-b-11-0-UI.s1 cDNA, 3' end /clon
477F4	3	405	AI394671	Hs.159678	0	2	tg24a07.x1 cDNA, 3' end /clone=IMAGE:2109684
472F5	194	366	NM_018490	Hs.160271	1.00E-93	1	G protein-coupled receptor 48 (GPR48), mRNA /
468B11	72	481	AI393041	Hs.160273	0	1	tg25b10.x1 cDNA, 3' end /clone=IMAGE:2109787
477D3	5	484	AI393906	Hs.160401	0	2	tg05f08.x1 cDNA, 3' end /clone=IMAGE:2107911
477D12	11	389	AI393962	Hs.160405	1.00E-178	1	tg11d08.x1 cDNA, 3' end /clone=IMAGE:2108463
477D5	15	262	AI393992	Hs.160408	1.00E-138	1	tg06c05.x1 cDNA, 3' end /clone=IMAGE:2107976
65A9	4106	5547	AF137030	Hs.160417	0	5	transmembrane protein 2 (TMEM2) mRNA, complete
513A2	4109	5547	NM_013390	Hs.160417	0	5	transmembrane protein 2 (TMEM2), mRNA /cds=(14
463F12	688	1425	AF218032	Hs.160422	0	1	clone PP902 unknown mRNA /cds=(693,1706) /gb=
165C1	2625	2987	X85116	Hs.160483	0	1	H.sapiens epb72 gene exon 1 /cds=(61,927) /gb=X85116 /gi=1
469G4	145	550	AI634652	Hs.160795	0	1	wa07e10.x1 cDNA, 3' end /clone=IMAGE:2297418
472C7	343	565	AI760020	Hs.160951	1.00E-105	1	wh83b05.x1 cDNA, 3' end /clone=IMAGE:2387313
466F12	485	662	BF207290	Hs.160954	2.00E-62	1	601870777F1 cDNA, 5' end /clone=IMAGE:4100850
477C10	5	290	BF437585	Hs.160980	1.00E-149	1	7p74d12.x1 cDNA, 3' end /clone=IMAGE:3651526
61E8	4435	6593	U83115	Hs.161002	0	3	non-lens beta gamma-crystallin like protein (AIM1) m
458E5	1	462	R84314	Hs.161043	1.00E-159	1	yq23a02.r1 cDNA, 5' end /clone=IMAGE:274443 /
466E12	117	447	BF001821	Hs.161075	0	1	7g93g02.x1 cDNA, 3' end /clone=IMAGE:3314066
102H4	7	219	AW963155	Hs.161786	1.00E-111	1	EST375228 /gb=AW963155 /gi=8152991 /ug=
118B6	2050	2260	NM_022570	Hs.161786	2.00E-75	1	C-type (calcium dependent, carbohydrate-reco
593C4	3863	4092	U86453	Hs.162808	9.00E-92	1	phosphatidylinositol 3-kinase catalytic subunit p1
467B7	129	455	AI023714	Hs.163442	1.00E-164	1	ow91h05.x1 cDNA, 3' end /clone=IMAGE:1654233
107G8	592	1016	AK023670	Hs.163495	0	1	FLJ13608 fis, clone PLACE1010628 /cds=UNK
74F3	229	449	AA627122	Hs.163787	4.00E-77	1	nq70g02.s1 cDNA, 3' end /clone=IMAGE:1157714
68B3	1094	1771	AK023494	Hs.164005	0	5	FLJ13432 fis, clone PLACE1002537 /cds=UNK
469H10	420	850	NM_002993	Hs.164021	0	1	small inducible cytokine subfamily B (Cys-X-C
464E9	86	424	AA811244	Hs.164168	1.00E-166	1	ob58h11.s1 cDNA, 3' end /clone=IMAGE:1335621
467E11	788	1330	NM_007063	Hs.164170	0	1	vascular Rab-GAP/TBC-containing (VRP), mRNA
597C5	59	1251	AY007135	Hs.164280	1.00E-126	3	clone CDABP0051 mRNA sequence /cds=(89,985) /
464H11	2	202	BF689700	Hs.164675	9.00E-65	1	602186609F1 cDNA, 5' end /clone=IMAGE:4298402
459D5	6	496	AI248204	Hs.165051	0	1	qh64h11.x1 cDNA, 3' end /clone=IMAGE:1849509
120F12	23	502	NM_001017	Hs.165590	1.00E-159	5	ribosomal protein S13 (RPS13), mRNA /cds=(32,4
469C11	301	613	AW364833	Hs.165681	1.00E-136	1	QV3-DT0043-211299-044-d03 cDNA /gb=AW364833
465D3	289	481	AI766638	Hs.165693	2.00E-62	1	wi02a10.x1 cDNA, 3' end /clone=IMAGE:2389050
465D6	107	238	AW850041	Hs.165695	3.00E-61	1	IL3-CT0216-170300-097-C07 cDNA /gb=AW850041
466C7	166	421	AI538546	Hs.165696	1.00E-122	1	td08b07.x1 cDNA, 3' end /clone=IMAGE:2075029
469C4	351	691	AI436561	Hs.165703	1.00E-148	1	ti03b03.x1 cDNA, 3' end /clone=IMAGE:2129357
62A12	32	256	AV727063	Hs.165980	1.00E-120	4	AV727063 cDNA, 5' end /clone=HTCCED11 /clone_
107C2	2427	2613	AJ250865	Hs.165986	1.00E-82	1	for TESS 2 protein (TESS /cds=(128,1393) /
461D5	1762	1935	NM_004031	Hs.166120	8.00E-81	1	interferon regulatory factor 7 (IRF7), transc

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

147D11	38	1283	AL022097	Hs.166203	0	5	DNA sequence from PAC 256G22 on chromosome 6p24
595H12	1321	1597	NM_002636	Hs.166204	1.00E-135	2	PHD finger protein 1 (PHF1), mRNA /cds=(56,1429
58H7	41	2036	AL136711	Hs.166254	0	2	mRNA; cDNA DKFZp566I133 (from clone DKFZp566I1
98D12	5559	6110	NM_014646	Hs.166318	0	1	lipin 2 (LPIN2), mRNA /cds=(239,2929) /gb=NM_0
468G1	146	509	AW873324	Hs.166338	1.00E-168	2	hl92a07.x1 cDNA, 3' end /clone=IMAGE:3009396
477D7	2900	3748	L14922	Hs.166563	0	1	DNA-binding protein (PO-GA) mRNA, complete cd
177E7	3265	3595	L23320	Hs.166563	0	1	replication factor C large subunit mRNA, complete cds
584H2	206	1613	NM_006925	Hs.166975	1.00E-112	5	splicing factor, arginine/serine-rich 5 (SFR
481F5	647	917	NM_002643	Hs.166982	1.00E-128	1	phosphatidylinositol glycan, class F (PIGF),
598E4	112	538	NM_002788	Hs.167106	1.00E-174	1	proteasome (prosome, macropain) subunit, alp
466D8	46	470	AI805131	Hs.167206	0	1	td11f04.x1 cDNA, 3' end /clone=IMAGE:2075359
464C8	342	469	BE674762	Hs.167208	4.00E-50	1	7e98d05.x1 cDNA, 3' end /clone=IMAGE:3293193
468A6	1177	1417	NM_003658	Hs.167218	4.00E-85	1	BarH-like homeobox 2 (BARX2), mRNA /cds=(96,93
74H10	1	1271	AF107405	Hs.167460	0	12	pre-mRNA splicing factor (SFRS3) mRNA, comple
60E9	3154	3926	U43185	Hs.167503	1.00E-143	2	signal transducer and activator of transcription Sta
517G3	1129	2787	NM_006994	Hs.167741	0	3	butyrophilin, subfamily 3, member A3 (BTN3A3),
175H2	2261	2467	U90548	Hs.167741	2.00E-86	1	butyrophilin (BTF3) mRNA, complete cds /cds=(171,192
588H5	1324	1735	NM_002901	Hs.167791	0	1	reticulocalbin 1, EF-hand calcium binding dom
331D7	53	625	AF116909	Hs.167827	4.00E-22	1	clone HH419 unknown mRNA /cds=(189,593) /gb=A
39C11	938	1672	AF026402	Hs.168103	0	1	U5 snRNP 100 kD protein mRNA, cds /cds=(39,2501
583C8	906	1669	NM_004818	Hs.168103	0	5	prp28, U5 snRNP 100 kd protein (U5-100K), mRNA
43B1	1156	1224	AF031167	Hs.168132	1.00E-22	1	interleukin 15 precursor (IL-15) mRNA, complet
479A7	424	801	NM_000585	Hs.168132	1.00E-149	1	interleukin 15 (IL15), mRNA /cds=(316,804) /g
67D6	1783	2336	AK024030	Hs.168232	0	1	FLJ13968 fis, clone Y79AA1001493, weakly
122H3	1646	2894	NM_023079	Hs.168232	0	2	hypothetical protein FLJ13855 (FLJ13855), mR
459H3	9	504	AI392830	Hs.168287	0	1	tg10b09.x1 cDNA, 3' end /clone=IMAGE:2108345
463G5	103	851	NM_003002	Hs.168289	0	1	succinate dehydrogenase complex, subunit D,
144G9	5588	5937	AL049935	Hs.168350	0	2	DKFZp564O1116 (from clone DKFZp564O
459A9	2293	2727	NM_000201	Hs.168383	0	2	intercellular adhesion molecule 1 (CD54), hum
123G3	2194	2675	AB046801	Hs.168640	0	2	mRNA for KIAA1581 protein, partial cds /cds=(0
112H10	505	864	AF007155	Hs.168694	1.00E-175	2	clone 23763 unknown mRNA, partial cds /cds=(0,
60H7	223	897	AF083420	Hs.168913	0	1	brain-specific STE20-like protein kinase 3 (
105C12	1698	2052	AK026671	Hs.169078	1.00E-176	1	FLJ23018 fis, clone LNG00903 /cds=(27,14
181B9	1148	1610	NM_003937	Hs.169139	0	1	kynureninase (L-kynurenine hydrolase) (KYNU)
462B7	13	478	AA977148	Hs.169168	0	1	oq24g08.s1 cDNA, 3' end /clone=IMAGE:1587326
41H5	197	624	U58913	Hs.169191	0	1	chemokine (hmrp-2a) mRNA, complete cds /cds=(71,484)
69G6	11	552	BF214508	Hs.169248	1.00E-160	4	601845758F1 cDNA, 5' end /clone=IMAGE:4076510
460B2	904	2904	NM_003202	Hs.169294	1.00E-161	2	transcription factor 7 (T-cell specific, HMG-
464G12	543	994	D26121	Hs.169303	0	1	mRNA for ZFM1 protein alternatively spliced product,
464B5	163	762	NM_013259	Hs.169330	0	1	neuronal protein (NP25), mRNA /cds=(49,897) /
593G4	787	1353	Z97989	Hs.169370	0	2	DNA sequence from PAC 66H14 on chromosome 6q21-22. Con
165F12	1177	1751	AK001725	Hs.169407	0	1	cDNA FLJ10863 fis, clone NT2RP4001575, highly
483B12	10871	11349	NM_004010	Hs.169470	0	1	dystrophin (muscular dystrophy, Duchenne and
518B3	22	1257	NM_002046	Hs.169476	0	5	glyceraldehyde-3-phosphate dehydrogenase (
67E7	1289	1597	U34995	Hs.169476	3.00E-88	1	normal keratinocyte substraction library mRNA, clon
47E9	2148	2452	NM_005461	Hs.169487	1.00E-172	1	Kreisler (mouse) maf-related leucine zipper h
69C3	846	3195	U41387	Hs.169531	0	24	Gu protein mRNA, partial cds /cds=(0,2405) /gb=U41387

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

468G7	73	450	AI523598	Hs.169541	1.00E-178	1	th08g11.x1 cDNA, 3' end /clone=IMAGE:2117732
72E12	490	3074	AJ251595	Hs.169610	0	29	for transmembrane glycoprotein (CD44 gen
471F2	97	533	AW172306	Hs.169738	0	1	xj37a08.x1 cDNA, 3' end /clone=IMAGE:2659382
589D4	96	488	NM_000994	Hs.169793	1.00E-163	2	ribosomal protein L32 (RPL32), mRNA /cds=(34,4
105B6	1590	2215	AK027212	Hs.169854	0	1	FLJ23559 fis, clone LNG09844 /cds=UNKNOW
462A8	1043	1529	NM_000305	Hs.169857	0	1	paraoxonase 2 (PON2), mRNA /cds=(32,1096) /gb
175D11	390	929	AF061736	Hs.169895	1.00E-132	2	ubiquitin-conjugating enzyme RIG-B mRNA, com
149A2	2442	2942	U75686	Hs.169900	0	1	polyadenylate binding protein mRNA, complete
524B9	2484	2709	NM_007049	Hs.169963	1.00E-125	2	butyrophilin, subfamily 2, member A1 (BTN2A1),
169G8	1192	1684	U90543	Hs.169963	0	1	butyrophilin (BTF1) mRNA, complete cds /cds=(210,179
129E9	686	1227	X70340	Hs.170009	0	1	transforming growth factor alpha /cds=(3
589C1	1893	3451	NM_004350	Hs.170019	0	5	runt-related transcription factor 3 (RUNX3),
331E1	5084	5496	NM_001621	Hs.170087	0	1	aryl hydrocarbon receptor (AHR) mRNA /cds=(643
595H7	659	4185	NM_002838	Hs.170121	0	34	protein tyrosine phosphatase, receptor type,
184G8	1083	3762	Y00062	Hs.170121	0	10	T200 leukocyte common antigen (CD45, LC-A) /c
109D4	4529	4876	AF032885	Hs.170133	0	1	forkhead protein (FKHR) mRNA, complete cds /cd
98A12	4529	4882	NM_002015	Hs.170133	1.00E-160	1	forkhead box O1A (rhabdomyosarcoma) (FOXO1A),
99E3	2098	2334	NM_004761	Hs.170160	1.00E-125	1	RAB2, member RAS oncogene family-like (RAB2L),
498F10	3472	4909	AL161952	Hs.170171	0	28	mRNA; cDNA DKFZp434M0813 (from clone DKFZp434M
465G7	390	462	AI475666	Hs.170288	2.00E-31	1	tc93c08.x1 cDNA, 3' end /clone=IMAGE:2073710
467E6	68	482	AK025743	Hs.170296	0	1	cDNA: FLJ22090 fis, clone HEP16084 /cds=UNKNOW
459H9	4659	5168	NM_014636	Hs.170307	0	1	Ral guanine nucleotide exchange factor RalGPS
38D9	618	992	D89678	Hs.170311	0	25	for A+U-rich element RNA binding factor,
589F11	1033	2022	NM_005463	Hs.170311	0	13	heterogeneous nuclear ribonucleoprotein D-l
469B9	127	573	AI436418	Hs.170326	0	1	ti01h02.x1 cDNA, 3' end /clone=IMAGE:2129235
183E4	2725	3777	NM_002444	Hs.170328	0	7	moesin (MSN), mRNA /cds=(100,1833) /gb=NM_002
170G2	1693	3305	Z98946	Hs.170328	0	4	DNA sequence from clone 376D21 on chromosome Xq11.1-12
464F6	162	534	AI492865	Hs.170331	1.00E-163	1	th78a05.x1 cDNA, 3' end /clone=IMAGE:2124752
472F8	412	554	AI373163	Hs.170333	1.00E-75	1	qz13a07.x1 cDNA, 3' end /clone=IMAGE:2021364
473C3	376	610	AW291507	Hs.170381	1.00E-123	1	UI-H-BI2-aga-g-11-0-UI.s1 cDNA, 3' end /clon
465E5	421	547	BE676049	Hs.170584	3.00E-54	1	7f21a03.x1 cDNA, 3' end /clone=IMAGE:3295276
477A3	25	202	AI475884	Hs.170587	4.00E-92	2	tc95c12.x1 cDNA, 3' end /clone=IMAGE:2073910
477A4	34	489	AI475905	Hs.170588	0	1	tc95f06.x1 cDNA, 3' end /clone=IMAGE:2073923
469F2	238	490	AI478556	Hs.170777	2.00E-84	1	tm53e03.x1 cDNA, 3' end /clone=IMAGE:2161852
472C5	357	474	AI479022	Hs.170784	1.00E-53	1	tm30a05.x1 cDNA, 3' end /clone=IMAGE:2158064
477D6	23	407	AI492034	Hs.170909	0	2	tg06f12.x1 cDNA, 3' end /clone=IMAGE:2108015
471D4	187	416	AI492181	Hs.170913	1.00E-106	1	tg07e06.x1 cDNA, 3' end /clone=IMAGE:2108098
464F8	14	142	AI492651	Hs.170934	7.00E-53	1	qz18b10.x1 cDNA, 3' end /clone=IMAGE:2021851
466D3	173	461	AI540204	Hs.170935	1.00E-131	1	td10h12.x1 cDNA, 3' end /clone=IMAGE:2075303
478F10	314	461	AI761144	Hs.171004	4.00E-45	1	wh97h01.x1 cDNA, 3' end /clone=IMAGE:2388721
476E2	187	253	AI494612	Hs.171009	2.00E-30	2	qz17a03.x1 cDNA, 3' end /clone=IMAGE:2021740
107G12	2413	2929	AK024436	Hs.171118	0	1	for FLJ00026 protein, partial cds /cds=(0
478H3	1237	1509	AL161725	Hs.171118	1.00E-107	1	DNA sequence from clone RP11-165F24 on chromosome 9.
477H10	252	489	BE674709	Hs.171120	3.00E-87	1	7e94f05.x1 cDNA, 3' end /clone=IMAGE:3292833
477H11	18	521	AI524202	Hs.171122	0	1	th10d11.x1 cDNA, 3' end /clone=IMAGE:2117877
466C10	24	216	BE816212	Hs.171261	8.00E-81	1	MR1-BN0212-280600-001-c06 cDNA /gb=BE816212
470A4	22	562	AI628893	Hs.171262	0	1	ty95h02.x1 cDNA, 3' end /clone=IMAGE:2286867
477C4	216	464	AI540161	Hs.171264	1.00E-112	2	td10c10.x1 cDNA, 3' end /clone=IMAGE:2075250
519E12	1	321	NM_016468	Hs.171566	1.00E-167	2	hypothetical protein (LOC51241), mRNA /cds=(
44C11	5363	5829	AF012872	Hs.171625	0	1	phosphatidylinositol 4-kinase 230 (pi4K230)
517D4	19	559	NM_003197	Hs.171626	0	3	transcription elongation factor B (SIII), pol

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

48E9	1563	1809	NM_004417	Hs.171695	1.00E-138	2	dual specificity phosphatase 1 (DUSP1), mRNA
520H5	941	3667	NM_002719	Hs.171734	0	2	protein phosphatase 2, regulatory subunit B (
106G2	1	308	BF243010	Hs.171774	1.00E-167	2	601877795F1 cDNA, 5' end /clone=IMAGE:4106303
524A7	14	359	NM_015933	Hs.171774	0	14	hypothetical protein (HSPC016), mRNA /cds=(3
117A11	311	614	BF966361	Hs.171802	1.00E-143	2	602286929F1 cDNA, 5' end /clone=IMAGE:4375783
38H11	885	2087	M55543	Hs.171862	0	6	guanylate binding protein isoform II (GBP-2) mRNA, co
512F8	232	1971	NM_004120	Hs.171862	0	12	guanylate binding protein 2, interferon-induc
111B9	3748	4161	NM_004941	Hs.171872	0	1	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide
192H11	5738	5903	NM_000937	Hs.171880	2.00E-68	1	polymerase (RNA) II (DNA directed) polypeptide
176F11	1322	4789	AL109935	Hs.171917	0	3	DNA sequence from clone RP5-1022P6 on chromosome 20 C
596G12	2472	3152	NM_001110	Hs.172028	0	5	a disintegrin and metalloproteinase domain 10
170A5	2438	2767	AK023154	Hs.172035	0	1	FLJ13092 fis, clone NT2RP3002147 /cds=(34
469D11	71	535	AI474074	Hs.172070	0	1	ti68h11.x1 cDNA, 3' end /clone=IMAGE:2137221
100G4	5574	5662	U02882	Hs.172081	3.00E-24	1	rolipram-sensitive 3',5'-cyclic AMP phosphodiester
524A11	1	2517	AL110202	Hs.172089	0	20	mRNA; cDNA DKFZp586I2022 (from clone DKFZp586
49A2	929	2845	NM_002568	Hs.172182	0	30	poly(A)-binding protein, cytoplasmic 1 (PABP
54C5	929	2484	Y00345	Hs.172182	0	9	polyA binding protein /cds=(502,2403) /gb=Y0
586B1	1042	1504	NM_002408	Hs.172195	0	1	mannosyl (alpha-1,6-)-glycoprotein beta-1,2
169H6	5576	5958	D25538	Hs.172199	0	1	KIAA0037 gene, complete cds /cds=(265,3507)
115G7	4531	4976	NM_001114	Hs.172199	0	1	adenylate cyclase 7 (ADCY7), mRNA /cds=(265,35
120F2	1	2496	NM_007363	Hs.172207	0	11	non-POU-domain-containing, octamer-binding
74A3	860	1364	Y11289	Hs.172207	0	1	p54nrb gene, exon 3 (and joined /cds=(136,1551)
60B7	695	1160	NM_000202	Hs.172458	0	1	iduronate 2-sulfatase (Hunter syndrome) (IDS
479D10	4059	4347	NM_000632	Hs.172631	1.00E-125	1	integrin, alpha M (complement component recep
167B10	1	389	NM_003761	Hs.172684	0	4	vesicle-associated membrane protein 8 (endob
189E11	1773	2038	NM_001345	Hs.172690	1.00E-149	2	diacylglycerol kinase, alpha (80kD) (DGKA), m
177C2	983	1489	X62535	Hs.172690	0	1	diacylglycerol kinase /cds=(103,2310)
458B12	535	1002	NM_012326	Hs.172740	0	1	microtubule-associated protein, RP/EB family
53A11	69	430	W26908	Hs.172762	1.00E-180	1	16b3 /gb=W26908 /gi=1306136 /ug=Hs.17276
151H2	2016	2572	M80359	Hs.172766	0	1	protein p78 mRNA, complete cds /cds=(171,2312) /gb=M8
100G10	3983	4302	AB037808	Hs.172789	1.00E-149	1	for KIAA1387 protein, partial cds /cds=(0
515D9	354	548	NM_004182	Hs.172791	3.00E-65	1	ubiquitously-expressed transcript (UXT), mR
193D9	2282	2757	AL109669	Hs.172803	0	3	mRNA full length insert cDNA clone EUROIMAGE 31
460H10	12	490	NM_016466	Hs.172918	0	1	hypothetical protein (LOC51239), mRNA /cds=(
483D3	3473	3941	AB011102	Hs.173081	0	1	mRNA for KIAA0530 protein, partial cds /cds=(0,
195B9	380	854	NM_005729	Hs.173125	0	2	peptidylprolyl isomerase F (cyclophilin F) (
173H6	6008	6412	NM_006283	Hs.173159	0	1	transforming, acidic coiled-coil containing
113E6	142	240	AI554733	Hs.173182	3.00E-49	1	tn27f08.x1 cDNA, 3' end /clone=IMAGE:2168871
56G8	140	630	AK002009	Hs.173203	0	2	FLJ11147 fis, clone PLACE1006678, weakly
69E6	1	463	BF131656	Hs.173205	1.00E-147	8	601820483F1 cDNA, 5' end /clone=IMAGE:4052348
44A2	6	196	X06347	Hs.173255	1.00E-94	1	U1 small nuclear RNP-specific A protein /cds=
149G1	79	498	AY007165	Hs.173274	1.00E-117	2	clone CDABP0163 mRNA sequence /cds=UNKNOWN /g
464F3	53	500	AW005376	Hs.173280	0	1	ws94a12.x1 cDNA, 3' end /clone=IMAGE:2505598
587H5	3299	4083	NM_014633	Hs.173288	0	2	KIAA0155 gene product (KIAA0155), mRNA /cds=(
499B9	1032	1923	NM_012081	Hs.173334	0	2	ELL-RELATED RNA POLYMERASE II, ELONGATION FAC
54F11	368	1923	U88629	Hs.173334	0	2	RNA polymerase II elongation factor ELL2, complete cd
459A4	2170	2775	AK001362	Hs.173374	0	1	cDNA FLJ10500 fis, clone NT2RP2000369 /cds=UNK

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

124B1	2566	3019	AB046825	Hs.173422	0	1	mRNA for KIAA1605 protein, partial cds /cds=(3
126H6	1080	1626	NM_006363	Hs.173497	0	1	Sec23 (S. cerevisiae) homolog B (SEC23B), mRNA
596D5	1233	1365	NM_004550	Hs.173611	8.00E-63	5	NADH dehydrogenase (ubiquinone) Fe-S protein
108C5	1709	1864	AK022681	Hs.173685	2.00E-83	1	FLJ12619 fis, clone NT2RM4001682 /cds=(39
583D12	3	1960	AK025703	Hs.173705	0	4	cDNA: FLJ22050 fis, clone HEP09454 /cds=UNKNOW
70B6	579	1140	AL049610	Hs.173714	0	2	DNA sequence from clone 1055C14 on chromosome Xq22.1-
46D7	590	1150	NM_012286	Hs.173714	0	1	MORF-related gene X (KIAA0026), mRNA /cds=(305
467G5	17	283	AA534537	Hs.173720	1.00E-104	1	nf80h10.s1 cDNA, 3' end /clone=IMAGE:926275 /
168H5	1	1066	D25274	Hs.173737	0	5	mRNA, clone:PO2ST9 /cds=UNKNOWN /gb=D25274 /
471B8	5347	5922	NM_014832	Hs.173802	0	1	KIAA0603 gene product (KIAA0603), mRNA /cds=(
177F4	1053	1622	U51166	Hs.173824	0	1	G/T mismatch-specific thymine DNA glycosylase mRNA,
471C3	396	719	AF277292	Hs.173840	1.00E-176	1	C4orf1 mRNA /cds=(0,281) /gb=AF277292 /gi=96
477F7	2053	2694	U80735	Hs.173854	0	3	CAGF28 mRNA, partial cds /cds=(0,2235) /gb=U80
41F3	3595	3890	M37435	Hs.173894	1.00E-143	1	macrophage-specific colony-stimulating factor (CSF
460C8	1542	1939	NM_014225	Hs.173902	0	1	protein phosphatase 2 (formerly 2A), regulator
458A9	292	414	AI763121	Hs.173904	4.00E-57	1	wi06d12.x1 cDNA, 3' end /clone=IMAGE:2389463
170B10	1230	3510	AL137681	Hs.173912	1.00E-176	5	cDNA DKFZp434M0326 (from clone DKFZp434M
126E10	1061	1795	Z17227	Hs.173936	1.00E-111	2	mRNA for transmebrane receptor protein /cds=(4
72H7	1210	1907	U08316	Hs.173965	0	2	insulin-stimulated protein kinase 1 (ISPK-1) mRNA, c
123G7	554	858	NM_005777	Hs.173993	1.00E-168	1	RNA binding motif protein 6 (RBM6), mRNA /cds=(
469C8	261	528	BE674902	Hs.174010	1.00E-113	1	7e97a04.x1 cDNA, 3' end /clone=IMAGE:3293070
117G6	2450	2657	NM_003089	Hs.174051	1.00E-112	1	small nuclear ribonucleoprotein 70kD polypept
103A5	4907	5011	NM_002209	Hs.174103	1.00E-48	1	integrin, alpha L (antigen CD11A (p180), lymph
159F4	333	925	AF261087	Hs.174131	0	7	DNA-binding protein TAXREB107 mRNA, complete
588F9	333	926	NM_000970	Hs.174131	0	8	ribosomal protein L6 (RPL6), mRNA /cds=(26,892
187A2	2993	3464	NM_001096	Hs.174140	0	2	ATP citrate lyase (ACLY), mRNA /cds=(84,3401)
41C6	3652	3992	X03663	Hs.174142	0	1	c-fms proto-oncogene /cds=(300,3218) /gb=X0
465G10	199	489	BE674951	Hs.174144	1.00E-152	1	7e97g10.x1 cDNA, 3' end /clone=IMAGE:3293154
468H10	28	159	AI524263	Hs.174193	6.00E-62	1	th11g07.x1 cDNA, 3' end /clone=IMAGE:2118012
99C7	402	733	NM_006435	Hs.174195	1.00E-179	2	interferon induced transmembrane protein 2 (
467E4	162	516	BF062628	Hs.174215	1.00E-157	1	7h62h05.x1 cDNA, 3' end /clone=IMAGE:3320601
74E5	2	485	D63789	Hs.174228	0	15	DNA for SCM-1beta precursor, complete cds /cd
470F11	108	305	AI590337	Hs.174258	1.00E-104	1	tn49c03.x1 cDNA, 3' end /clone=IMAGE:2171716
463D2	1	194	AV734916	Hs.175971	1.00E-94	1	AV734916 cDNA, 5' end /clone=cdAAHE11 /clone_
477E5	75	222	AI380955	Hs.176374	2.00E-33	1	tg18b08.x1 cDNA, 3' end /clone=IMAGE:2109111
473A9	1	296	AI708327	Hs.176430	1.00E-162	1	at04c02.x1 cDNA, 3' end /clone=IMAGE:2354114
468C3	24	235	AW081098	Hs.176498	6.00E-91	1	xc29a12.x1 cDNA, 3' end /clone=IMAGE:2585662
479D11	595	1810	J04162	Hs.176663	0	14	leukocyte IgG receptor (Fc-gamma-R) mRNA, complete c
108G2	388	579	AI638800	Hs.176920	6.00E-78	4	tt32e01.x1 cDNA, 3' end /clone=IMAGE:2242488
467A10	98	170	AI865603	Hs.177045	6.00E-27	1	wk47g03.x1 cDNA, 3' end /clone=IMAGE:2418580
117A6	1179	1403	AF116606	Hs.177415	1.00E-112	2	PRO0890 mRNA, complete cds /cds=(1020,1265) /
73F2	236	919	NM_016406	Hs.177507	0	4	hypothetical protein (HSPC155), mRNA /cds=(2
516D8	24	340	NM_006886	Hs.177530	1.00E-179	1	ATP synthase, H+ transporting, mitochondrial
479F4	163	676	NM_002414	Hs.177543	0	1	antigen identified by monoclonal antibodies 1
126A9	906	2105	NM_005534	Hs.177559	0	35	interferon gamma receptor 2 (interferon gamma
41H6	905	1826	U05875	Hs.177559	0	10	clone pSK1 interferon gamma receptor accessory factor
37G1	1690	2420	U62961	Hs.177584	0	1	succinyl CoA:3-oxoacid CoA transferase precursor (O
597H7	1764	2520	AF218002	Hs.177596	0	7	clone PP2464 unknown mRNA /cds=(675,2339) /gb

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

520B8	1036	1202	NM_006888	Hs.177656	4.00E-90	3	calmodulin 1 (phosphorylase kinase, delta) (C
151G7	2439	3048	J03473	Hs.177766	0	1	poly(ADP-ribose) synthetase mRNA, complete cds /cds=
116C6	318	834	BC001980	Hs.177781	1.00E-144	4	clone MGC:5618, mRNA, complete cds /cds=(156,
179C11	211	737	X07834	Hs.177781	0	3	manganese superoxide dismutase (EC 1.15.1.1)
98A9	213	648	M73547	Hs.178112	0	4	polyposis locus (DP1 gene) mRNA, complete cds /cds=(82
459E10	149	789	AK023719	Hs.178357	0	1	cDNA FLJ13657 fis, clone PLACE1011563 /cds=(8
120H6	137	404	NM_021029	Hs.178391	1.00E-136	1	ribosomal protein L44 (RPL44), mRNA /cds=(37,3
589E9	371	596	NM_000973	Hs.178551	1.00E-125	1	ribosomal protein L8 (RPL8), mRNA /cds=(43,816
142F5	1848	2210	D21090	Hs.178658	1.00E-179	1	XP-C repair complementing protein (p58/HHR23
120H11	402	532	AV716627	Hs.178703	9.00E-69	1	AV716627 cDNA, 5' end /clone=DCBBCH05 /clone_
98G11	3287	6017	NM_004859	Hs.178710	0	5	clathrin, heavy polypeptide (Hc) (CLTC), mRNA
177H1	142	421	BF130300	Hs.178732	1.00E-139	1	601818357F1 cDNA, 5' end /clone=IMAGE:4041902
472A10	421	562	AI681868	Hs.178784	4.00E-63	1	tx50a12.x1 cDNA, 3' end /clone=IMAGE:2272990
467G6	194	292	AW138461	Hs.179003	1.00E-49	1	UI-H-BI1-adg-e-06-0-UI.s1 cDNA, 3' end /clon
465C11	3312	3606	NM_016562	Hs.179152	1.00E-166	1	toll-like receptor 7 (LOC51284), mRNA /cds=(13
469F7	268	405	AI568459	Hs.179419	3.00E-45	1	tn39e07.x1 cDNA, 3' end /clone=IMAGE:2170020
99F11	750	2687	NM_006472	Hs.179526	0	73	upregulated by 1,25-dihydroxyvitamin D-3 (VD
39G9	526	2687	S73591	Hs.179526	0	17	brain-expressed HHCPA78 homolog VDUP1 (Gene)
102A1	2235	2659	AL034343	Hs.179565	0	1	DNA sequence from clone RP1-108C2 on chromosome 6p12.
492B2	1074	2126	NM_002717	Hs.179574	1.00E-131	3	protein phosphatase 2 (formerly 2A), regulator
143F2	242	457	NM_005771	Hs.179608	1.00E-117	1	retinol dehydrogenase homolog (RDHL]) mRNA /
111G7	626	898	NM_002659	Hs.179657	1.00E-153	1	plasminogen activator, urokinase receptor (P
585D2	61	3189	AL162068	Hs.179662	0	6	mRNA; cDNA DKFZp762G106 (from clone DKFZp762G1
125G4	1159	1627	NM_000389	Hs.179665	1.00E-130	2	cyclin-dependent kinase inhibitor 1A (p21, Ci
331A1	51	377	AK026642	Hs.179666	1.00E-161	2	FLJ22989 fis, clone KAT11824, highly sim
516H12	19	362	NM_000997	Hs.179779	1.00E-180	3	ribosomal protein L37 (RPL37), mRNA /cds=(28,3
170A11	1390	2087	L20298	Hs.179881	0	1	transcription factor (CBFB) mRNA, 3' end /cds=(
195H8	1732	2110	NM_001755	Hs.179881	1.00E-173	1	core-binding factor, beta subunit (CBFB), tra
127G6	2406	2924	AK022499	Hs.179882	0	2	cDNA FLJ12437 fis, clone NT2RM1000118, weakly
461E6	610	1148	NM_014153	Hs.179898	0	1	HSPC055 protein (HSPC055), mRNA /cds=(1400,19
516B3	4	584	NM_000975	Hs.179943	1.00E-136	2	ribosomal protein L11 (RPL11), mRNA /cds=(0,53
62F8	24	537	X79234	Hs.179943	1.00E-175	1	ribosomal protein L11 /cds=(0,536) /gb=
471B11	1990	2496	NM_005802	Hs.179982	0	1	tumor protein p53-binding protein (TP53BPL),
194B4	693	956	NM_004159	Hs.180062	1.00E-112	1	proteasome (prosome, macropain) subunit, bet
49D4	1002	1259	NM_002690	Hs.180107	1.00E-125	1	polymerase (DNA directed), beta (POLB), mRNA
184A11	26	515	AK024823	Hs.180139	0	2	FLJ21170 fis, clone CAS10946, highly sim
593A8	43	535	NM_006937	Hs.180139	0	13	SMT3 (suppressor of mif two 3, yeast) homolog 2
61D10	102	722	AF161415	Hs.180145	0	1	HSPC297 mRNA, partial cds /cds=(0,438) /gb=AF
178A4	131	628	NM_017924	Hs.180201	0	2	hypothetical protein FLJ20671 (FLJ20671), mR
463H9	54	171	NM_005507	Hs.180370	1.00E-60	1	cofilin 1 (non-muscle) (CFL1), mRNA /cds=(51,5
162B9	2139	2386	AB013382	Hs.180383	1.00E-124	1	for DUSP6, complete cds /cds=(351,1496) /
190B7	1743	2386	NM_001946	Hs.180383	1.00E-124	2	dual specificity phosphatase 6 (DUSP6), trans
589B11	21	1566	NM_006597	Hs.180414	0	11	heat shock 70kD protein 8 (HSPA8), mRNA /cds=(8
73G2	21	1567	Y00371	Hs.180414	0	16	hsc70 gene for 71 kd heat shock protein /cds=(83,2023)
62G1	985	1559	X89602	Hs.180433	0	1	rTS beta protein /cds=(17,1267) /gb=X896
98F9	1479	3653	L38951	Hs.180446	0	9	importin beta subunit mRNA, complete cds /cds=(
590F12	283	614	NM_001026	Hs.180450	0	1	ribosomal protein S24 (RPS24), mRNA /cds=(142,
597F2	2670	3046	AF187554	Hs.180532	0	47	sperm antigen-36 mRNA, complete cds /cds=(234,
482E2	85	366	AL571386	Hs.180546	1.00E-106	1	AL571386 cDNA /clone=CS0DI009YL09-(3-prime)

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

109C2	324	682	BE540238	Hs.180549	1.00E-143	1	601059809F1 cDNA, 5' end /clone=IMAGE:3446283
68G8	1447	3594	AF123094	Hs.180566	0	3	API2-MLT fusion protein (API2-MLT) mRNA, comp
180B9	1851	2142	NM_002087	Hs.180577	1.00E-160	2	granulin (GRN), mRNA /cds=(62,1843) /gb=NM_00
51E4	880	2466	NM_005066	Hs.180610	0	6	splicing factor proline/glutamine rich (poly
50G4	880	1280	X70944	Hs.180610	0	1	PTB-associated splicing factor /cds=(85
127C8	317	3175	AK023143	Hs.180638	0	5	cDNA FLJ13081 fis, clone NT2RP3002033 /cds=(17
125E2	287	1692	AL117621	Hs.180777	0	2	mRNA; cDNA DKFZp564M0264 (from clone DKFZp564
521F11	1969	2431	AF126964	Hs.180799	0	1	C3HC4-type zinc finger protein (LZK1) mRNA, co
479C11	1186	2245	AK000271	Hs.180804	1.00E-155	2	cDNA FLJ20264 fis, clone COLF7912 /cds=UNKNOWN
479C2	732	911	NM_001242	Hs.180841	3.00E-62	1	tumor necrosis factor receptor superfamily, m
596D2	67	942	NM_000977	Hs.180842	0	11	ribosomal protein L13 (RPL13), mRNA /cds=(51,6
41E9	884	1779	AL050337	Hs.180866	0	2	DNA sequence from clone 503F13 on chromosome 6q24.1-25
196C10	679	1338	NM_000416	Hs.180866	0	2	interferon gamma receptor 1 (IFNGR1), mRNA /cd
99A10	1	1655	AF218029	Hs.180877	0	11	clone PP781 unknown mRNA /cds=(113,523) /gb=A
65H9	1	1320	Z48950	Hs.180877	0	6	hH3.3B gene for histone H3.3 /cds=(10,420) /gb=Z
160G1	2065	2538	AF045555	Hs.180900	0	2	wbscr1 (WBSCR1) and wbscr5 (WBSCR5) genes, com
596B1	5	860	NM_001008	Hs.180911	0	5	ribosomal protein S4, Y-linked (RPS4Y), mRNA
192F11	1857	2521	AK000299	Hs.180952	0	1	cDNA FLJ20292 fis, clone HEP05374 /cds=(21,140
75D10	94	1656	AY007118	Hs.181013	0	8	clone CDABP0006 mRNA sequence /cds=(20,784) /
46H2	105	1661	NM_002629	Hs.181013	0	5	phosphoglycerate mutase 1 (brain) (PGAM1), mR
107G10	4869	5527	AK024391	Hs.181043	0	1	FLJ14329 fis, clone PLACE4000259, highly
179A1	22	908	AK001934	Hs.181112	0	2	FLJ11072 fis, clone PLACE1004982 /cds=(2
118D5	610	1130	NM_014166	Hs.181112	0	1	HSPC126 protein (HSPC126), mRNA /cds=(25,837)
483D9	659	915	X57809	Hs.181125	1.00E-123	1	rearranged immunoglobulin lambda light chain mRNA /c
596B10	499	1198	NM_005517	Hs.181163	0	2	high-mobility group (nonhistone chromosomal)
74A12	34	1674	AK026650	Hs.181165	0	192	FLJ22997 fis, clone KAT11962, highly sim
99H8	1079	2742	BC001412	Hs.181165	0	260	eukaryotic translation elongation factor 1
70F10	144	840	AB015798	Hs.181195	0	1	HSJ2 mRNA for DnaJ homolog, complete cds /cds=
64E10	72	856	BC002446	Hs.181195	0	2	MRJ gene for a member of protein family, clone
597F6	1119	1767	NM_001675	Hs.181243	0	3	activating transcription factor 4 (tax-respon
109D8	825	1233	D32129	Hs.181244	0	1	HLA class-I (HLA-A26) heavy chain, complete c
593H10	465	1222	NM_016057	Hs.181271	0	3	CGI-120 protein (LOC51644), mRNA /cds=(37,570
127H10	4782	5154	AB020335	Hs.181300	0	1	Pancreas-specific TSA305 mRNA , complete cds
150F7	509	1238	M11353	Hs.181307	1.00E-175	5	H3.3 histone class C mRNA, complete cds /cds=(374,784)
127F7	895	1057	NM_002107	Hs.181307	3.00E-85	2	H3 histone, family 3A (H3F3A), mRNA /cds=(374,7
39H10	6	416	BF676042	Hs.181357	0	7	602084011F1 cDNA, 5' end /clone=IMAGE:4248195
99G12	193	842	NM_002295	Hs.181357	0	28	laminin receptor 1 (67kD, ribosomal protein SA
66A12	312	1084	M20430	Hs.181366	0	4	MHC class II HLA-DR-beta (DR2-DQw1/DR4 DQw3) mRNA, co
71H11	748	1096	NM_002125	Hs.181366	1.00E-176	1	major histocompatibility complex, class II,
56E4	272	521	AI827911	Hs.181400	1.00E-126	1	wf34e11.x1 cDNA, 3' end /clone=IMAGE:2357516
170F6	5255	5724	D63486	Hs.181418	0	1	KIAA0152 gene, complete cds /cds=(128,1006)
464A11	5981	6322	NM_014730	Hs.181418	1.00E-159	1	KIAA0152 gene product (KIAA0152), mRNA /cds=(
514F6	1	232	AW955745	Hs.181426	1.00E-117	1	EST367815 cDNA /gb=AW955745 /gi=8145428 /ug=
177E2	690	947	U81002	Hs.181466	1.00E-130	2	TRAF4 associated factor 1 mRNA, partial cds /c
99B5	260	1660	NM_001549	Hs.181874	0	6	interferon-induced protein with tetratricope

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

595H9	104	645	M90356	Hs.181967	0	1	BTF3 protein homologue gene, complete cds /cds=(0,644
67E2	1057	1782	AK026664	Hs.182225	4.00E-85	3	FLJ23011 fis, clone LNG00572 /cds=(288,7
190A3	319	1615	NM_014052	Hs.182238	0	7	GW128 protein (GW128), mRNA /cds=(698,889) /g
140B10	1770	2034	U46751	Hs.182248	2.00E-92	1	phosphotyrosine independent ligand p62 for the Lck S
158H11	371	597	D50420	Hs.182255	1.00E-126	1	OTK27, complete cds /cds=(94,480) /gb
584A12	95	1397	NM_005008	Hs.182255	0	3	non-histone chromosome protein 2 (S. cerevisia
40G2	735	908	Y00503	Hs.182265	7.00E-41	1	keratin 19 /cds=(32,1234) /gb=Y00503 /gi=340
596E7	1	886	NM_001743	Hs.182278	0	3	calmodulin 2 (phosphorylase kinase, delta) (C
129E10	36	350	L29348	Hs.182378	1.00E-174	2	granulocyte-macrophage colony-stimulating
487G1	184	934	NM_002952	Hs.182426	0	3	ribosomal protein S2 (RPS2), mRNA /cds=(240,90
517G6	126	1497	NM_005742	Hs.182429	0	4	protein disulfide isomerase-related protein
60E12	10	1329	M16342	Hs.182447	0	4	nuclear ribonucleoprotein particle (hnRNP) C protein
98E9	10	1184	NM_004500	Hs.182447	0	8	heterogeneous nuclear ribonucleoprotein C (
496A4	87	1835	NM_014394	Hs.182470	0	2	PTD010 protein (PTD010), mRNA /cds=(129,1088)
110F11	947	1571	AF061738	Hs.182579	0	2	leucine aminopeptidase mRNA, complete cds /cd
124E1	1330	1889	NM_005739	Hs.182591	0	2	RAS guanyl releasing protein 1 (calcium and DA
143B2	32	565	Z47087	Hs.182643	0	1	RNA polymerase II elongation factor-like
103D2	161	538	NM_001015	Hs.182740	8.00E-97	5	ribosomal protein S11 (RPS11), mRNA /cds=(15,4
331C2	1310	1585	D64015	Hs.182741	1.00E-136	1	for T-cluster binding protein, complete c
59E9	27	269	BF245224	Hs.182825	1.00E-105	1	601863885F1 cDNA, 5' end /clone=IMAGE:4082396
525E3	12	261	NM_007209	Hs.182825	1.00E-135	2	ribosomal protein L35 (RPL35), mRNA /cds=(27,3
70C9	189	625	BE963551	Hs.182928	1.00E-129	1	601657346R1 cDNA, 3' end /clone=IMAGE:3866266
177B9	14	561	BF242969	Hs.182937	0	2	601877739F1 cDNA, 5' end /clone=IMAGE:4106289
519H3	34	526	NM_021130	Hs.182937	0	1	peptidylprolyl isomerase A (cyclophilin A) (
159A5	3163	3579	AK026491	Hs.182979	1.00E-141	2	FLJ22838 fis, clone KAIA4494, highly sim
106G11	2956	3527	AF204231	Hs.182982	1.00E-138	2	88-kDa Golgi protein (GM88) mRNA, complete cds
169A3	2117	2495	M33336	Hs.183037	1.00E-105	3	cAMP-dependent protein kinase type I-alpha subunit (
124H9	2767	2955	NM_002734	Hs.183037	7.00E-91	1	protein kinase, cAMP-dependent, regulatory,
107B3	2877	3182	U17989	Hs.183105	1.00E-170	1	nuclear autoantigen GS2NA mRNA, complete cds /
476A6	538	893	NM_016523	Hs.183125	0	1	killer cell lectin-like receptor F1 (KLRF1), m
75A1	629	1222	AK001433	Hs.183297	0	1	FLJ10571 fis, clone NT2RP2003121, weakly
597E11	97	1656	AF248966	Hs.183434	0	5	HT028 mRNA, complete cds /cds=(107,1159) /gb=
124A2	2015	2756	AK024275	Hs.183506	0	1	cDNA FLJ14213 fis, clone NT2RP3003572 /cds=(11
74F2	2082	2418	U53347	Hs.183556	1.00E-177	2	neutral amino acid transporter B mRNA, complete cds /
482C5	1211	1688	NM_018399	Hs.183656	0	1	VNN3 protein (HSA238982), mRNA /cds=(45,1550)
594H12	1718	3458	NM_001418	Hs.183684	0	4	eukaryotic translation initiation factor 4 g
61H11	1457	2024	U73824	Hs.183684	0	2	p97 mRNA, complete cds /cds=(306,3029) /gb=U73824 /g
75H7	342	2258	M26880	Hs.183704	0	7	ubiquitin mRNA, complete cds /cds=(135,2192) /gb=M26
599E7	2306	3111	D44640	Hs.183706	0	6	HUMSUPY040 cDNA /clone=035-00-1 /gb=D44640 /
518H4	1554	1973	NM_002078	Hs.183773	0	1	golgi autoantigen, golgin subfamily a, 4 (GOL
520C3	98	255	NM_018955	Hs.183842	3.00E-64	1	ubiquitin B (UBB), mRNA /cds=(94,783) /gb=NM_
102C11	1730	1808	M15182	Hs.183868	8.00E-33	2	beta-glucuronidase mRNA, complete cds /cds=(26,1981
523D3	1730	2183	NM_000181	Hs.183868	0	2	glucuronidase, beta (GUSB), mRNA /cds=(26,198
187A12	122	828	NM_003589	Hs.183874	0	1	cullin 4A (CUL4A), mRNA /cds=(160,2139) /gb=N
156F4	228	907	AF119665	Hs.184011	0	4	inorganic pyrophosphatase complete cds
525B8	225	791	NM_021129	Hs.184011	0	2	pyrophosphatase (inorganic) (PP), nuclear ge
589B1	3	394	NM_000993	Hs.184014	0	10	ribosomal protein L31 (RPL31), mRNA /cds=(7,38

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

99D6	3909	4308	NM_004985	Hs.184050	1.00E-145	1	v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene
166B3	12	345	BE964596	Hs.184052	1.00E-90	1	601658521R1 cDNA, 3' end /clone=IMAGE:3885796
591G6	1348	1958	NM_022152	Hs.184052	0	3	PP1201 protein (PP1201), mRNA /cds=(66,1001)
114E11	1780	1942	AK025645	Hs.184062	4.00E-59	1	cDNA: FLJ21992 fis, clone HEP06554 /cds=(60,84
597E4	8	407	NM_000982	Hs.184108	1.00E-114	6	ribosomal protein L21 (gene or pseudogene) (RP
162C5	295	1062	L41887	Hs.184167	0	3	splicing factor, arginine/serine-rich 7 (SFR
109F6	151	749	AF054182	Hs.184211	0	1	mitochondrial processing peptidase beta-subu
462C6	4590	5087	NM_015001	Hs.184245	0	1	KIAA0929 protein Msx2 interacting nuclear tar
517D1	1510	1936	NM_004252	Hs.184276	1.00E-162	7	solute carrier family 9 (sodium/hydrogen exch
55E3	174	427	NM_018370	Hs.184465	1.00E-107	1	hypothetical protein FLJ11259 (FLJ11259), mR
50F9	2484	3108	AB023182	Hs.184523	0	1	for KIAA0965 protein, partial cds /cds=(0
100A4	297	1941	AK025730	Hs.184542	1.00E-149	3	FLJ22077 fis, clone HEP12728, highly sim
113D4	950	1623	NM_016061	Hs.184542	0	1	CGI-127 protein (LOC51646), mRNA /cds=(125,49
145D11	41	339	BE730026	Hs.184582	1.00E-111	1	601562642F1 cDNA, 5' end /clone=IMAGE:3832258
595F4	69	548	NM_000986	Hs.184582	0	1	ribosomal protein L24 (RPL24), mRNA /cds=(39,5
108H10	250	701	U00946	Hs.184592	0	1	clone A9A2BRB5 (CAC)n/(GTG)n repeat-containing mRN
43B5	4399	4488	AF104032	Hs.184601	3.00E-24	1	L-type amino acid transporter subunit LAT1 mRN
104F12	298	1713	NM_014999	Hs.184627	0	2	KIAA0118 protein (KIAA0118), mRNA /cds=(255,9
122E8	513	995	AF035307	Hs.184697	0	2	clone 23785 mRNA sequence /cds=UNKNOWN /gb=AF
40H2	66	2605	M37197	Hs.184760	1.00E-177	4	CCAAT-box-binding factor (CBF) mRNA, complete cds /c
514E4	29	519	NM_000984	Hs.184776	0	3	ribosomal protein L23a (RPL23A), mRNA /cds=(2
589A7	736	983	AK025533	Hs.184793	1.00E-138	1	cDNA: FLJ21880 fis, clone HEP02743 /cds=UNKNOW
142G5	1918	2157	AL049782	Hs.184938	8.00E-83	3	Novel human gene mapping to chomosome 13 /cds=UNKNOWN /gb=A
462G9	178	398	AI085568	Hs.185062	1.00E-76	1	oy68b05.x1 cDNA, 3' end /clone=IMAGE:1670961
470C12	81	333	T98171	Hs.185575	1.00E-105	1	ye56c12.s1 cDNA, 3' end /clone=IMAGE:121750 /
463F2	3175	3359	NM_014686	Hs.186840	1.00E-72	1	KIAA0355 gene product (KIAA0355), mRNA /cds=(
461E4	907	1118	NM_018519	Hs.186874	4.00E-91	1	hypothetical protein PRO2266 (PRO2266), mRNA
155A1	53	379	AI619574	Hs.187362	1.00E-109	1	ty50c09.x1 cDNA, 3' end /clone=IMAGE:2282512
461C9	2948	3458	NM_014504	Hs.187660	0	1	putative Rab5 GDP/GTP exchange factor homologu
470F2	5	331	BE646499	Hs.187872	1.00E-156	1	7e87h02.x1 cDNA, 3' end /clone=IMAGE:3292179
68D12	590	740	AW963239	Hs.187908	4.00E-66	1	EST375312 /gb=AW963239 /gi=8153075 /ug=
75H12	2012	2585	AL110269	Hs.187991	0	1	cDNA DKFZp564A122 (from clone DKFZp564A1
167G4	1474	1958	NM_015626	Hs.187991	0	1	DKFZP564A122 protein (DKFZP564A122), mRNA /c
137G3	54	197	AI625368	Hs.188365	2.00E-34	46	ts37c10.x1 cDNA, 3' end /clone=IMAGE:2230770
464C12	183	404	AA432364	Hs.188777	7.00E-94	1	zw76a09.s1 cDNA, 3' end /clone=IMAGE:782104 /
467E9	29	183	AA576947	Hs.188886	1.00E-63	1	nm82b04.s1 cDNA, 3' end /clone=IMAGE:1074703
467B4	349	459	AI392805	Hs.189031	2.00E-49	1	tg04h03.x1 cDNA, 3' end /clone=IMAGE:2107829
461E2	242	473	BE674964	Hs.190065	1.00E-109	1	7f11b09.x1 cDNA, 3' end /clone=IMAGE:3294329
466F4	58	295	BG326781	Hs.190219	1.00E-132	1	602425659F1 cDNA, 5' end /clone=IMAGE:4563471
465H4	111	558	AA582958	Hs.190229	0	1	nn80d08.s1 cDNA, 3' end /clone=IMAGE:1090191
470F9	26	529	AI763206	Hs.190453	0	1	wh95e09.x1 cDNA, 3' end /clone=IMAGE:2388520
66H12	1	3459	D00099	Hs.190703	0	5	for Na,K-ATPase alpha-subunit, complete
472E1	338	540	AW294083	Hs.190904	2.00E-46	1	UI-H-BI2-ahg-b-05-0-UI.s1 cDNA, 3' end /clon
522G10	433	970	NM_003757	Hs.192023	0	2	eukaryotic translation initiation factor 3,
54G8	29	410	AW838827	Hs.192123	0	1	CM1-LT0059-280100-108-e02 /gb=AW838827
465G4	261	515	BF224348	Hs.192463	1.00E-104	1	7q86c05.x1 cDNA /clone=IMAGE /gb=BF224348 /g
468F9	392	487	AI524039	Hs.192524	2.00E-36	1	tg99h02.x1 cDNA, 3' end /clone=IMAGE:2116947
466C6	111	392	AW972048	Hs.192534	1.00E-153	1	EST384032 cDNA /gb=AW972048 /gi=8161789 /ug=

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

184F12	1	677	AF090927	Hs.192705	0	1	clone HQ0457 PRO0457 mRNA, complete cds /cds=(
464C11	1	65	BE298181	Hs.192755	3.00E-23	1	601118566F1 cDNA, 5' end /clone=IMAGE:3028193
465H3	108	706	BG036938	Hs.192965	0	1	602287708F1 cDNA, 5' end /clone=IMAGE:4375153
169F9	4138	4890	D87454	Hs.192966	0	1	KIAA0265 gene, partial cds /cds=(0,1205) /gb
118H10	1104	1858	AK024263	Hs.193063	1.00E-132	2	cDNA FLJ14201 fis, clone NT2RP3002955 /cds=UNK
472F3	28	405	BF062295	Hs.193237	0	1	7k76b11.x1 cDNA, 3' end /clone=IMAGE:3481293
40A5	1933	2611	X12830	Hs.193400	0	1	interleukin-6 (IL-6) receptor /cds=(437,184
63B5	327	582	AW959162	Hs.193669	1.00E-103	1	EST371232 /gb=AW959162 /gi=8148846 /ug=
52G10	803	1173	M57627	Hs.193717	0	1	interleukin 10 (IL10) mRNA, complete cds /cds=(30,566
469F5	2088	2438	AL110204	Hs.193784	1.00E-179	1	mRNA; cDNA DKFZp586K1922 (from clone DKFZp586K
598H7	1428	1715	NM_014828	Hs.194035	1.00E-119	1	KIAA0737 gene product (KIAA0737), mRNA /cds=(
462B6	103	546	BE618004	Hs.194362	1.00E-165	1	601462354F1 cDNA, 5' end /clone=IMAGE:3865861
472F12	1177	1667	AB036737	Hs.194369	0	2	mRNA for RERE, complete cds /cds=(636,5336) /g
182E10	11785	13486	U82828	Hs.194382	0	5	ataxia telangiectasia (ATM) gene, complete cd
458F4	258	408	NM_022739	Hs.194477	2.00E-62	1	E3 ubiquitin ligase SMURF2 (SMURF2), mRNA /cd
583D2	1425	1732	NM_014232	Hs.194534	1.00E-136	1	vesicle-associated membrane protein 2 (synapt
38H8	1198	1620	U89387	Hs.194638	0	1	RNA polymerase II subunit hsRPB4 gene, complete cds /
122H10	5292	5481	NM_023005	Hs.194688	4.00E-80	1	bromodomain adjacent to zinc finger domain, 1B
186G9	1	1908	AL136945	Hs.194718	0	2	mRNA; cDNA DKFZp586O012 (from clone DKFZp586O0
113F3	1852	2375	NM_000634	Hs.194778	0	1	interleukin 8 receptor, alpha (IL8RA), mRNA /
106A3	35	404	U11870	Hs.194778	0	1	interleukin-8 receptor type A (IL8RBA) gene, promote
473B8	1001	1314	AF319438	Hs.194976	1.00E-172	1	SH2 domain-containing phosphatase anchor pro
57F9	442	1934	Y14039	Hs.195175	0	27	mRNA for CASH alpha protein /cds=(481,1923) /g
49E5	2314	2512	NM_018666	Hs.195292	2.00E-37	1	putative tumor antigen (SAGE), mRNA /cds=(167,
473B10	406	532	BE671815	Hs.195374	1.00E-54	1	7a47c12.x1 cDNA, 3' end /clone=IMAGE:3221878
595B5	59	311	AI653766	Hs.195378	6.00E-46	1	ty01b06.x1 cDNA, 3' end /clone=IMAGE:2277779
60G4	42	1554	D13642	Hs.195614	0	2	KIAA0017 gene, complete cds /cds=(136,1335)
473B9	739	927	AF241534	Hs.196015	2.00E-73	1	hydatidiform mole associated and imprinted (H
99C10	1075	1424	NM_000294	Hs.196177	1.00E-115	1	phosphorylase kinase, gamma 2 (testis) (PHKG2
45H9	956	1405	AF283645	Hs.196270	0	1	folate transporter/carrier mRNA, complete cd
54F9	2567	2954	U04636	Hs.196384	0	1	cyclooxygenase-2 (hCox-2) gene, complete cds /cds=(1
38F12	401	606	AI984074	Hs.196398	1.00E-104	1	wz56c02.x1 cDNA, 3' end /clone=IMAGE:2562050
157G1	403	551	AJ006835	Hs.196769	7.00E-77	2	RNA transcript from U17 small nucleolar RNA ho
163F4	1	402	AI650871	Hs.197028	0	1	wa95f03.x1 cDNA, 3' end /clone=IMAGE:2303933
160B3	408	476	AI832038	Hs.197091	5.00E-27	1	wj99e02.x1 3' end /clone=IMAGE:2410970
105E8	1299	3674	AB020657	Hs.197298	0	6	for KIAA0850 protein, complete cds /cds=(
178G12	2097	3593	AF205218	Hs.197298	0	8	NS1-binding protein-like protein mRNA, compl
585F1	284	1711	NM_001469	Hs.197345	0	4	thyroid autoantigen 70kD (Ku antigen) (G22P1)
39C10	545	1984	Z83840	Hs.197345	0	2	DNA sequence from clone CTA-216E10 on chromosome 22 C
58E12	2162	3013	NM_001530	Hs.197540	0	2	hypoxia-inducible factor 1, alpha subunit (ba
125G11	3673	4059	D29805	Hs.198248	0	1	mRNA for beta-1,4-galactosyltransferase, complete
41H10	6	821	M33906	Hs.198253	1.00E-156	2	MHC class II HLA-DQA1 mRNA, complete cds /cds=(43,810)
186A11	551	1031	NM_004544	Hs.198271	0	2	NADH dehydrogenase (ubiquinone) 1 alpha subco
126D8	993	1381	NM_021105	Hs.198282	0	1	phospholipid scramblase 1 (PLSCR1), mRNA /cds
174C12	4824	5257	NM_003070	Hs.198296	0	1	SWI/SNF related, matrix associated, actin dep

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

109C6	128	833	X04327	Hs.198365	0	1	erythrocyte 2,3-bisphosphoglycerate mutase mRNA EC
64B12	4383	5289	NM_000189	Hs.198427	0	2	hexokinase 2 (HK2), mRNA /cds=(1490,4243) /gb
70B4	3267	5289	Z46376	Hs.198427	0	4	HK2 mRNA for hexokinase II /cds=(1490,4243) /gb=Z
478H6	186	475	AI978581	Hs.198694	1.00E-129	1	wq72d08.x1 cDNA, 3' end /clone=IMAGE:2476815
587G1	767	1143	NM_006837	Hs.198767	1.00E-170	1	COP9 (constitutive photomorphogenic, Arabido
465F12	373	554	BE621611	Hs.198802	2.00E-77	1	601493754T1 cDNA, 3' end /clone=IMAGE:3895836
123B3	310	3608	AB011108	Hs.198891	0	3	mRNA for KIAA0536 protein, partial cds /cds=(0,
157H3	3457	5268	D50929	Hs.198899	0	2	KIAA0139 gene, complete cds /cds=(128,4276)
477H1	35	592	NM_002229	Hs.198951	0	1	jun B proto-oncogene (JUNB), mRNA /cds=(253,12
53C5	979	1296	X51345	Hs.198951	1.00E-160	1	jun-B mRNA for JUN-B protein /cds=(253,1296) /gb=X513
54H8	350	501	AW450874	Hs.199014	5.00E-81	1	UI-H-BI3-all-a-11-0-UI.s1 cDNA, 3' end /clon
520E12	3506	3878	L04731	Hs.199160	0	1	translocation T(4:11) of ALL-1 gene to chromoso
57F4	5941	6266	NM_006267	Hs.199179	1.00E-158	1	RAN binding protein 2 (RANBP2), mRNA /cds=(127,
50B10	5	3645	D86984	Hs.199243	0	2	KIAA0231 gene, partial cds /cds=(0,1430) /gb
68E12	1757	2052	L25124	Hs.199248	1.00E-156	2	prostaglandin E2 receptor mRNA, complete cds /
484H3	1879	1958	NM_000958	Hs.199248	3.00E-33	1	prostaglandin E receptor 4 (subtype EP4) (PTGE
466G6	368	3287	NM_013233	Hs.199263	0	2	Ste-20 related kinase (SPAK), mRNA /cds=(173,1
464B9	633	1068	AF015041	Hs.199291	0	1	NUMB-R protein (NUMB-R) mRNA, complete cds /c
522F9	2	116	AI669591	Hs.200442	5.00E-59	1	tw34b09.x1 cDNA, 3' end /clone=IMAGE:2261561
60F11	4945	5114	AB040942	Hs.201500	7.00E-92	1	for KIAA1509 protein, partial cds /cds=(0
72D12	819	1293	AF104398	Hs.201673	0	1	cornichon mRNA, complete cds /cds=(56,490) /g
105G5	1629	2130	AF091263	Hs.201675	0	1	RNA binding motif protein 5 (RBM5) mRNA, comple
116G3	1637	2854	NM_005778	Hs.201675	0	2	RNA binding motif protein 5 (RBM5), mRNA /cds=(
40A10	254	431	AI693179	Hs.201789	5.00E-85	1	wd68d12.x1 cDNA, 3' end /clone=IMAGE:2336759
473D4	421	547	BE551203	Hs.201792	3.00E-49	1	7b55h12.x1 cDNA, 3' end /clone=IMAGE:3232199
472D8	313	623	AW390251	Hs.202402	1.00E-123	1	CM4-ST0182-051099-021-b06 cDNA /gb=AW390251
66H5	176	482	AI271437	Hs.203041	1.00E-173	1	qi19c05.x1 cDNA, 3' end /clone=IMAGE:1856936
594C2	35	368	AW131782	Hs.203606	1.00E-147	2	xf34e08.x1 cDNA, 3' end /clone=IMAGE:2619974
138B12	101	420	AW194379	Hs.203755	1.00E-93	3	xm08h07.x1 3' end /clone=IMAGE:2683645
473D3	1	234	AI538474	Hs.203784	1.00E-117	1	td06h08.x1 cDNA, 3' end /clone=IMAGE:2074911
471A5	113	442	AI393908	Hs.203829	1.00E-153	1	tg05f10.x1 cDNA, 3' end /clone=IMAGE:2107915
40A4	1621	2037	AF004230	Hs.204040	0	1	monocyte/macrophage Ig-related receptor MIR
463H1	7	319	AW977671	Hs.204214	1.00E-161	1	EST389900 cDNA /gb=AW977671 /gi=8169049 /ug=
478E7	25	434	AI762023	Hs.204610	0	2	wh89f04.x1 cDNA, 3' end /clone=IMAGE:2387935
55E11	324	469	AI741246	Hs.204656	1.00E-58	12	wg26g09.x1 cDNA, 3' end /clone=IMAGE:2366272
478G10	345	476	AI760901	Hs.204703	9.00E-34	1	wi09h06.x1 cDNA, 3' end /clone=IMAGE:2389787
470E11	374	507	AI762741	Hs.204707	2.00E-49	1	wh93h02.x1 cDNA, 3' end /clone=IMAGE:2388339
478F5	179	437	AI086035	Hs.204873	1.00E-110	1	oy70h04.x1 cDNA, 3' end /clone=IMAGE:1671223
464G4	33	320	AI749444	Hs.204929	5.00E-50	1	at24c03.x1 cDNA, 3' end /clone=IMAGE:2356036
472D2	88	198	AI760018	Hs.205071	4.00E-54	1	wh83b02.x1 cDNA, 3' end /clone=IMAGE:2387307
470D9	5	422	AW976641	Hs.205079	0	1	EST388750 cDNA /gb=AW976641 /gi=8167872 /ug=
470D4	122	500	AA885473	Hs.205175	0	1	am10c12.s1 cDNA, 3' end /clone=IMAGE:1466422
473C5	285	525	BF679831	Hs.205319	2.00E-96	1	602154415F1 cDNA, 5' end /clone=IMAGE:4295595
470E7	295	521	AI762557	Hs.205327	9.00E-95	2	wh92f07.x1 cDNA, 3' end /clone=IMAGE:2388229
478F11	11	447	AI761141	Hs.205452	0	3	wh97g08.x1 cDNA, 3' end /clone=IMAGE:2388734
459A12	111	323	N72600	Hs.205555	9.00E-96	1	za46f08.s1 cDNA, 3' end /clone=IMAGE:295623 /
470F4	214	481	AW977820	Hs.205675	1.00E-131	1	EST389824 cDNA /gb=AW977820 /gi=8168971 /ug=
102G3	1	249	BF680988	Hs.205696	2.00E-78	1	602156272F1 cDNA, 5' end /clone=IMAGE:4297216
472B2	312	700	BF794256	Hs.206761	0	1	602255454F1 cDNA, 5' end /clone=IMAGE:4338949

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

470C1	1113	1643	AK024118	Hs.206868	0	1	cDNA FLJ14056 fis, clone HEMBB1000335 /cds=UNK
469H7	1076	1215	U15177	Hs.206984	3.00E-69	1	cosmid CRI-JC2015 at D10S289 in 10sp13 /cds=(0,1214)
61F9	5	181	AW340421	Hs.207995	4.00E-94	1	hc96h02.x1 cDNA, 3' end /clone=IMAGE:2907891
473C2	239	551	BF439675	Hs.208854	1.00E-151	1	nab69e11.x1 cDNA /clone=IMAGE /gb=BF439675 /
62G11	159	292	BE781611	Hs.208985	1.00E-60	1	601467463F1 cDNA, 5' end /clone=IMAGE:3870902
472E2	258	554	AI343473	Hs.209203	1.00E-135	1	tb97a08.x1 cDNA, 3' end /clone=IMAGE:2062262
471C10	148	498	AI768880	Hs.209511	0	1	wh71e04.x1 cDNA, 3' end /clone=IMAGE:2386206
470G9	416	561	AI798144	Hs.209609	4.00E-63	1	wh81g12.x1 cDNA, 3' end /clone=IMAGE:2387206
478C10	120	447	AI809310	Hs.210385	1.00E-158	2	wh75h08.x1 cDNA, 3' end /clone=IMAGE:2386623
476B7	64	341	AI075288	Hs.210727	1.00E-151	2	oy69h10.x1 cDNA, 3' end /clone=IMAGE:1671139
477G4	915	1541	AB040919	Hs.210958	0	1	mRNA for KIAA1486 protein, partial cds /cds=(1
468C2	215	498	AI832182	Hs.210995	1.00E-145	1	td13h11.x1 cDNA, 3' end /clone=IMAGE:2075589
472D11	1	300	AI860120	Hs.211024	1.00E-126	1	wh39e01.x1 cDNA, 3' end /clone=IMAGE:2383128
470D3	30	317	AW362304	Hs.211194	1.00E-137	1	CM3-CT0275-031199-031-a08 cDNA /gb=AW362304
179F6	105	551	AI823649	Hs.211535	0	1	wi85g03.x1 3' end /clone=IMAGE:2400148
477G12	2439	4050	NM_020993	Hs.211563	0	4	B-cell CLL/lymphoma 7A (BCL7A), mRNA /cds=(953
39A11	5178	5792	L10717	Hs.211576	0	2	T cell-specific tyrosine kinase mRNA, complete
187B9	5365	5790	NM_005546	Hs.211576	0	1	IL2-inducible T-cell kinase (ITK), mRNA /cds=
152C2	3965	4297	Z22551	Hs.211577	1.00E-174	1	kinectin gene /cds=(69,4139) /gb=Z22551 /gi=296
120A2	2556	2917	NM_005955	Hs.211581	0	1	metal-regulatory transcription factor 1 (MTF
147A2	2915	4407	M59465	Hs.211600	0	6	tumor necrosis factor alpha inducible protein A20 mRN
583B12	2404	3981	NM_006290	Hs.211600	0	11	tumor necrosis factor, alpha-induced protein
589F3	1905	2274	AF090693	Hs.211610	0	1	apoptosis-related RNA binding protein (NAPOR-
470G11	277	462	AI862623	Hs.211744	5.00E-99	1	wh99h10.x1 cDNA, 3' end /clone=IMAGE:2388931
473F2	195	423	BE675092	Hs.211828	2.00E-95	1	7f02d07.x1 cDNA, 3' end /clone=IMAGE:3293485
517D2	1059	1366	BC000747	Hs.211973	1.00E-162	2	Similar to homolog of Yeast RRP4 (ribosomal RN
109D9	391	533	AI922921	Hs.212553	2.00E-68	1	wn81c05.x1 cDNA, 3' end /clone=IMAGE:2452232
494H12	172	549	AI912585	Hs.213385	0	3	we11d07.x1 cDNA, 3' end /clone=IMAGE:2340781
596G11	4740	5687	AB007916	Hs.214646	0	8	mRNA for KIAA0447 protein, partial cds /cds=(2
104C12	843	1787	AL031282	Hs.215595	0	2	DNA sequence from clone 283E3 on chromosome 1p36.21-36
124F8	1391	2913	NM_002074	Hs.215595	0	4	guanine nucleotide binding protein (G protein)
157E8	1264	1627	AK001548	Hs.215766	0	4	FLJ10686 fis, clone NT2RP3000252, highly
519G3	1729	2094	NM_012341	Hs.215766	0	1	GTP-binding protein (NGB), mRNA /cds=(23,1924
473E7	2278	2472	AB022663	Hs.215857	3.00E-52	1	HFB30 mRNA, complete cds /cds=(236,1660) /gb=
104F7	4	1324	D00017	Hs.217493	0	3	for lipocortin II, complete cds /cds=(49,1
58G2	11	1324	NM_004039	Hs.217493	0	7	annexin A2 (ANXA2), mRNA /cds=(49,1068) /gb=N
467D4	27	443	AI392814	Hs.221014	1.00E-180	1	tg10a02.x1 cDNA, 3' end /clone=IMAGE:2108330
463B1	69	457	AV686223	Hs.221642	0	1	AV686223 cDNA, 5' end /clone=GKCGXH11 /clone_
464D10	295	552	BF058398	Hs.221695	1.00E-115	1	7k30d01.x1 cDNA, 3' end /clone=IMAGE:3476785
466C12	1	427	AI540165	Hs.222186	0	1	td10d05.x1 cDNA, 3' end /clone=IMAGE:2075241
125H10	2596	2917	AB046830	Hs.222746	0	1	mRNA for KIAA1610 protein, partial cds /cds=(0
473C4	1	193	BF435098	Hs.222833	9.00E-72	1	7p05g01.x1 cDNA, 3' end /clone=IMAGE:3645097
37B4	18	371	AW389509	Hs.223747	1.00E-147	1	CM3-ST0163-051099-019-b11 /gb=AW389509
470H7	106	357	AI766706	Hs.223935	1.00E-116	1	wi02g11.x1 cDNA, 3' end /clone=IMAGE:2389124
472D12	1	370	AL133721	Hs.224680	0	1	DKFZp761H09121_r1 cDNA, 5' end /clone=DKFZp76
124E4	53	208	AI874107	Hs.224760	7.00E-50	3	wm49b01.x1 cDNA, 3' end /clone=IMAGE:2439241
477G3	146	412	AI400714	Hs.225567	1.00E-141	1	tg93g12.x1 cDNA, 3' end /clone=IMAGE:2116390
112F12	2313	2799	AL163279	Hs.225674	0	1	chromosome 21 segment HS21C079 /cds=(0,6888)
118D12	6187	6775	NM_015384	Hs.225767	0	1	IDN3 protein (IDN3), mRNA /cds=(706,7182) /gb
109B7	2208	3315	AF119417	Hs.225939	0	2	nonfunctional GM3 synthase mRNA, alternativel

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

125A8	2877	3381	NM_006999	Hs.225951	0	1	topoisomerase-related function protein 4-1
129C8	5510	5893	AF012108	Hs.225977	0	1	Amplified in Breast Cancer (AIB1) mRNA, comple
39G12	4498	4859	NM_014977	Hs.227133	1.00E-93	2	KIAA0670 protein/acinus (KIAA0670), mRNA /cd
153D10	1	286	AF000145	Hs.227400	1.00E-139	2	germinal center kinase related protein kinase
464B12	901	1425	AL050131	Hs.227429	0	1	mRNA; cDNA DKFZp586I111 (from clone DKFZp586I1
459D9	3828	4314	NM_004841	Hs.227806	0	1	ras GTPase activating protein-like (NGAP), mR
135E9	135	773	NM_004049	Hs.227817	0	1	BCL2-related protein A1 (BCL2A1), mRNA /cds=(
59F10	123	808	Y09397	Hs.227817	0	12	GRS protein /cds=(102,629) /gb=Y09397 /
516H4	1901	2462	NM_014287	Hs.227823	0	1	pM5 protein (PM5), mRNA /cds=(0,3668) /gb=NM_0
107C12	2776	3390	Y15906	Hs.227913	0	1	for XAGL protein /cds=(132,1646) /gb=Y159
152C7	171	1390	AF052155	Hs.227949	0	2	clone 24761 mRNA sequence /cds=UNKNOWN /gb=AF
522G8	108	293	AI917348	Hs.228486	2.00E-70	1	ts83d10.x1 cDNA, 3' end /clone=IMAGE:2237875
66C7	304	445	AI094726	Hs.228795	1.00E-26	1	qa08f05.x1 cDNA, 3' end /clone=IMAGE:1686177
585D1	51	294	AI199388	Hs.228817	5.00E-73	1	qs75e05.x1 cDNA, 3' end /clone=IMAGE:1943936
468E9	113	324	AI523873	Hs.228926	7.00E-77	2	tg97c12.x1 cDNA, 3' end /clone=IMAGE:2116726
466F1	44	139	AI380491	Hs.229374	3.00E-39	2	tf95b10.x1 cDNA, 3' end /clone=IMAGE:2107003
182F1	40	465	AI354231	Hs.229385	1.00E-138	4	qv12c04.x1 cDNA, 3' end /clone=IMAGE:1981350
465C1	237	316	AW812896	Hs.229868	3.00E-38	1	RC3-ST0186-250200-018-a11 cDNA /gb=AW812896
178H7	42	353	AI581732	Hs.229918	1.00E-68	5	ar74f03.x1 cDNA, 3' end /clone=IMAGE:2128349
72H6	48	534	AI818777	Hs.229990	1.00E-85	3	wl11f10.x1 cDNA, 3' end /clone=IMAGE:2424619
181E9	52	279	AI827451	Hs.229993	1.00E-66	1	wl17d11.x1 cDNA, 3' end /clone=IMAGE:2425173
38H1	225	311	AI579979	Hs.230430	1.00E-25	1	tq45a01.x1 cDNA, 3' end /clone=IMAGE:2211720
489G11	66	369	AI818596	Hs.230492	1.00E-112	5	wk74d04.x1 cDNA, 3' end /clone=IMAGE:2421127
118D6	40	161	AI025427	Hs.230752	6.00E-37	1	ow27g06.s1 cDNA, 3' end /clone=IMAGE:1648090
462H6	305	437	AI087055	Hs.230805	3.00E-67	1	oy70c09.x1 cDNA, 3' end /clone=IMAGE:1671184
107C11	93	240	AI796419	Hs.230939	1.00E-40	1	wj17f02.x1 cDNA, 3' end /clone=IMAGE:2403099
591A1	65	316	AA767883	Hs.231154	7.00E-59	4	oa30h07.s1 cDNA, 3' end /clone=IMAGE:1306525
471B3	177	519	BE407125	Hs.231510	1.00E-166	1	601301818F1 cDNA, 5' end /clone=IMAGE:3636412
64G11	609	950	AL542592	Hs.231816	1.00E-166	1	AL542592 cDNA /clone=CS0DE012YA05-(5-prime)
108G1	1	210	AW006867	Hs.231987	1.00E-109	1	ws15d07.x1 cDNA, 3' end /clone=IMAGE:2497261
115F3	44	185	AW016002	Hs.232000	7.00E-75	2	UI-H-BI0p-abh-h-06-0-UI.s1 cDNA, 3' end /clo
138A6	4771	5194	D15050	Hs.232068	0	1	transcription factor AREB6, complete cds /cd
472A6	311	497	BF195579	Hs.232257	1.00E-78	1	7n85c03.x1 cDNA, 3' end /clone=IMAGE:3571205
111A7	285	463	AW026667	Hs.233261	1.00E-41	1	wv15d09.x1 cDNA, 3' end /clone=IMAGE:2529617
67G8	292	560	BE719483	Hs.233383	4.00E-94	3	MR1-HT0858-020800-001-c06 /gb=BE719483
123B11	180	351	AW006045	Hs.233560	5.00E-82	1	wz81b09.x1 cDNA, 3' end /clone=IMAGE:2565209
472E3	1	319	AW027530	Hs.233564	1.00E-180	1	wv74c06.x1 cDNA, 3' end /clone=IMAGE:2535274
36F11	943	1896	Z85996	Hs.233750	0	6	DNA sequence from PAC 431A14 on chromosome 6p21. Conta
184G6	49	491	BF694761	Hs.233936	0	9	602080851F2 cDNA, 5' end /clone=IMAGE:4245133
599C7	12	540	NM_006471	Hs.233936	0	55	myosin, light polypeptide, regulatory, non-s
156B4	405	774	AF054185	Hs.233952	1.00E-164	1	proteasome subunit HSPC complete cds /c
595G5	85	315	NM_002792	Hs.233952	1.00E-126	1	proteasome (prosome, macropain) subunit, alp
67F5	108	556	AK000654	Hs.234149	0	1	FLJ20647 fis, clone KAT02147 /cds=(90,836
591B6	1	555	NM_017918	Hs.234149	0	6	hypothetical protein FLJ20647 (FLJ20647), mR
111B7	1887	2217	AK023204	Hs.234265	1.00E-120	1	cDNA FLJ13142 fis, clone NT2RP3003212, modera
72F6	314	2581	AL035071	Hs.234279	0	2	DNA sequence from clone 1085F17 on chromosome 20q11.1
514H4	2105	2523	NM_012325	Hs.234279	0	1	microtubule-associated protein, RP/EB family
599A10	1	1163	NM_002300	Hs.234489	0	30	lactate dehydrogenase B (LDHB), mRNA /cds=(84
163A8	470	1153	X13794	Hs.234489	0	4	lactate dehydrogenase B gene exon 1 and (EC 1.1.1.
125E5	31	465	NM_000978	Hs.234518	1.00E-117	2	ribosomal protein L23 (RPL23), mRNA /cds=(25,4

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

471B1	1499	2033	L05148	Hs.234569	0	1	protein tyrosine kinase related mRNA sequence /cds=UN
466D7	1050	1402	NM_013451	Hs.234680	0	1	fer-1 (C.elegans)-like 3 (myoferlin) (FER1L3)
108B11	407	742	X14008	Hs.234734	0	1	lysozyme gene (EC 3.2.1.17) /cds=(82,474) /gb=X14008
476A12	3	440	AI076222	Hs.235042	0	2	oy65b09.x1 cDNA, 3' end /clone=IMAGE:1670681
464H7	994	2425	AL157426	Hs.235390	1.00E-22	1	mRNA; cDNA DKFZp761B101 (from clone DKFZp761B1
472F2	2203	2431	AK024137	Hs.235498	7.00E-97	1	cDNA FLJ14075 fis, clone HEMBB1001905, weakly
63C7	1159	1751	AK000260	Hs.235712	0	1	FLJ20253 fis, clone COLF6895 /cds=UNKNOWN
73C8	39	485	AI379474	Hs.235823	0	1	tc57g08.x1 cDNA, 3' end /clone=IMAGE:2068766
590H8	182	449	AA020845	Hs.235883	1.00E-145	3	ze64a07.r1 cDNA, 5' end /clone=IMAGE:363732 /
182H3	468	2009	NM_001535	Hs.235887	1.00E-119	5	HMT1 (hnRNP methyltransferase, S. cerevisiae)
119B12	253	596	NM_003075	Hs.236030	0	1	SWI/SNF related, matrix associated, actin dep
461C5	654	1112	AK026410	Hs.236449	0	1	cDNA: FLJ22757 fis, clone KAIA0803 /cds=(92,24
182G3	514	2817	AK023223	Hs.236494	0	2	FLJ13161 fis, clone NT2RP3003589, highly
469G7	857	1336	AK026359	Hs.236744	0	1	cDNA: FLJ22706 fis, clone HSI13163 /cds=UNKNOW
592A9	1522	1888	NM_020135	Hs.236828	0	1	putative helicase RUVBL (LOC56897), mRNA /cds
177A1	1260	1704	AK001514	Hs.236844	1.00E-170	1	FLJ10652 fis, clone NT2RP2005886 /cds=(50
594G2	916	1537	NM_018169	Hs.236844	0	2	hypothetical protein FLJ10652 (FLJ10652), mR
98D10	1881	1964	NM_006947	Hs.237825	9.00E-36	1	signal recognition particle 72kD (SRP72), mRN
72C7	36	1214	M29696	Hs.237868	0	2	interleukin-7 receptor (IL-7) mRNA, complete cds /cd
591B10	577	1658	NM_002185	Hs.237868	0	9	interleukin 7 receptor (IL7R), mRNA /cds=(22,1
109G2	16	405	AF116682	Hs.238205	0	1	PRO2013 mRNA, complete cds /cds=(135,380) /gb
41E1	2163	2733	U60805	Hs.238648	0	1	oncostatin-M specific receptor beta subunit (OSMRB)
599C11	508	1734	AK026110	Hs.238707	0	5	cDNA: FLJ22457 fis, clone HRC09925 /cds=(56,14
143E8	2	595	AV700542	Hs.238730	1.00E-177	6	AV700542 cDNA, 3' end /clone=GKCAFD05 /clone_
596C11	77	658	AW955090	Hs.238954	0	5	EST367160 cDNA /gb=AW955090 /gi=8144773 /ug=
169C7	1371	1634	AY004255	Hs.238990	1.00E-148	1	cdk inhibitor p27KIP1 mRNA, complete cds /cds=
173C1	1599	1859	BC001971	Hs.238990	1.00E-146	1	Similar to cyclin-dependent kinase inhibitor
458B5	1539	1809	AL136828	Hs.238996	1.00E-131	1	mRNA; cDNA DKFZp434K0427 (from clone DKFZp434K
591H9	6104	6559	AL157902	Hs.239114	0	1	DNA sequence from clone RP4-675C20 on chromosome 1p13
512G4	231	2376	NM_005746	Hs.239138	0	61	pre-B-cell colony-enhancing factor (PBEF), m
53D11	935	2053	U02020	Hs.239138	0	15	pre-B cell enhancing factor (PBEF) mRNA, complete cds
38B7	2187	2263	AK025021	Hs.239189	1.00E-36	1	FLJ21368 fis, clone COL03056, highly sim
458E10	90	622	NM_016533	Hs.239208	0	1	ninjurin 2 (NINJ2), mRNA /cds=(56,484) /gb=NM
184G10	1608	2056	AK026535	Hs.239307	0	1	FLJ22882 fis, clone KAT03587, highly sim
194D9	1544	1683	NM_003680	Hs.239307	4.00E-57	1	tyrosyl-tRNA synthetase (YARS), mRNA /cds=(0,
110C7	450	1216	AF246221	Hs.239625	0	4	transmembrane protein BRI mRNA, complete cds
599G9	446	1205	NM_021999	Hs.239625	0	13	integral membrane protein 2B (ITM2B), mRNA /cd
515E4	1404	1671	NM_014515	Hs.239720	1.00E-132	1	CCR4-NOT transcription complex, subunit 2 (C
115H10	1124	2079	BC000105	Hs.239760	0	2	Similar to CG14740 gene product, clone MGC:25
466E3	605	923	NM_005301	Hs.239891	1.00E-164	2	G protein-coupled receptor 35 (GPR35), mRNA /
52B5	993	1243	AJ223075	Hs.239894	1.00E-106	1	for TRIP protein /cds=(178,2532) /gb=AJ22
171E10	88	399	AW002624	Hs.240077	1.00E-145	1	wu60d10.x1 cDNA, 3' end /clone=IMAGE:990854 /
75C5	325	1604	AK027191	Hs.240443	0	8	FLJ23538 fis, clone LNG08010, highly sim
597D3	1134	1792	BC001255	Hs.240770	0	1	nuclear cap binding protein subunit 2, 20kD,
98A11	596	6834	NM_005385	Hs.241493	0	10	natural killer-tumor recognition sequence (N
98C10	1580	2204	AK027187	Hs.241507	0	40	cDNA: FLJ23534 fis, clone LNG06974, highly sim
463E8	324	846	AF047002	Hs.241520	0	1	transcriptional coactivator ALY mRNA, partia
514G6	802	1238	NM_012392	Hs.241531	0	3	peflin (PEF), mRNA /cds=(12,866) /gb=NM_01239

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

177G4	1375	1887	AF099149	Hs.241558	0	1	TRIAD1 type I mRNA, complete cds /cds=(144,1625
110E4	1320	1937	AK021704	Hs.241567	0	1	FLJ11642 fis, clone HEMBA1004356, highly
513B12	700	1447	NM_016839	Hs.241567	0	3	RNA binding motif, single stranded interacting
500G10	910	1249	NM_000594	Hs.241570	0	1	tumor necrosis factor (TNF superfamily, membe
514B6	735	1032	NM_018630	Hs.241576	1.00E-155	1	hypothetical protein PRO2577 (PRO2577), mRNA
590H9	61	251	NM_016200	Hs.241578	1.00E-104	1	U6 snRNA-associated Sm-like protein LSm8 (LOC
50A6	200	311	AK026704	Hs.242868	3.00E-57	3	FLJ23051 fis, clone LNG02642 /cds=UNKNOW
104C10	199	353	AA424812	Hs.243029	2.00E-74	1	zw04b02.s1 cDNA, 3' end /clone=IMAGE:768267 /
72G4	182	415	AW081232	Hs.243321	1.00E-99	4	xc22e08.x1 cDNA, 3' end /clone=IMAGE:2585030
521D12	32	287	AW102836	Hs.243457	6.00E-96	1	xd38h12.x1 cDNA, 3' end /clone=IMAGE:2596103
102F3	79	157	W45562	Hs.243720	4.00E-26	1	zc26e07.s1 cDNA, 3' end /clone=IMAGE:323460 /
56D6	193	454	M97856	Hs.243886	1.00E-122	1	histone-binding protein mRNA, complete cds /c
595D8	25	495	NM_002482	Hs.243886	0	1	nuclear autoantigenic sperm protein (histone-
46G5	2137	2661	AK000745	Hs.243901	0	1	cDNA FLJ20738 fis, clone HEP08257 /cds=UNKNOWN
477D4	141	250	AI394001	Hs.244666	4.00E-51	1	tg06d04.x1 cDNA, 3' end /clone=IMAGE:2107975
139B7	50	235	AW078847	Hs.244816	4.00E-32	2	xb18g07.x1 cDNA, 3' end /clone=IMAGE:2576700
472C4	74	464	AW139918	Hs.245138	0	1	UI-H-BI1-aee-d-05-0-UI.s1 cDNA, 3' end /clon
459F7	45	229	AW080951	Hs.245616	7.00E-58	1	xc28c10.x1 cDNA, 3' end /clone=IMAGE:2585586
100A6	41	1795	L22009	Hs.245710	1.00E-143	3	hnRNP H mRNA, complete cds /cds=(72,1421) /gb=L22009
592G8	41	1798	NM_005520	Hs.245710	0	6	heterogeneous nuclear ribonucleoprotein H1
71G4	382	583	AL136607	Hs.245798	1.00E-104	1	mRNA; cDNA DKFZp564I0422 (from clone DKFZp564
118B9	4495	5528	AK024391	Hs.246112	0	4	cDNA FLJ14329 fis, clone PLACE4000259, highly
471E5	148	464	AI568725	Hs.246299	1.00E-177	1	th15a01.x1 cDNA, 3' end /clone=IMAGE:2118312
464D11	26	526	N28843	Hs.246358	0	1	yx59d10.r1 cDNA, 5' end /clone=IMAGE:266035 /
40H7	550	1108	S57235	Hs.246381	0	1	CD68=110kda transmembrane glycoprotein [human, promonocy
471E12	152	507	AW117189	Hs.246494	1.00E-149	1	xd83f08.x1 cDNA, 3' end /clone=IMAGE:2604231
479C1	47	345	AV739961	Hs.246796	1.00E-140	1	AV739961 cDNA, 5' end /clone=CBFBRA10 /clone_
472C9	43	400	BF796642	Hs.246818	0	1	602259846F1 cDNA, 5' end /clone=IMAGE:4343171
47F11	2	227	AB015856	Hs.247433	1.00E-123	1	for ATF6, complete cds /cds=(68,2080) /gb
179H9	12	379	AL031313	Hs.247783	1.00E-111	1	DNA sequence from clone 581F12 on chromosome Xq21. Co
167A9	5	352	Z00013	Hs.247792	1.00E-163	5	H.sapiens germline gene for the leader peptide and variable
72B8	402	672	L15006	Hs.247824	1.00E-139	2	Ig superfamily CTLA-4 mRNA, complete cds /cds=
488H10	135	672	NM_005214	Hs.247824	1.00E-146	5	cytotoxic T-lymphocyte-associated protein 4
188G8	1	255	NM_002991	Hs.247838	1.00E-135	1	small inducible cytokine subfamily A (Cys-Cys
153D11	401	720	AL049545	Hs.247877	1.00E-133	2	DNA sequence from clone 263J7 on chromosome 6q14.3-15
44D2	42	448	AL035604	Hs.247894	1.00E-133	1	DNA sequence from clone 38C16 on chromosome 6q22.33-2
180B7	10	271	L21961	Hs.247947	4.00E-72	1	Ig rearranged lambda-chain mRNA, subgroup VL3, V-J re
110B11	311	803	U08626	Hs.247984	0	1	glutamine synthetase pseudogene /cds=(0,899) /gb=U
74G5	361	965	X14798	Hs.248109	0	1	DNA for c-ets-1 proto-oncogene /cds=(278,1603) /gb=
60H10	214	527	AW150084	Hs.248657	1.00E-99	3	xg36f03.x1 cDNA, 3' end /clone=IMAGE:2629661
64E2	329	536	BF512500	Hs.248689	1.00E-112	1	UI-H-BI3-alw-h-10-0-UI.s1 cDNA, 3' end /clon
470C6	278	470	AI832183	Hs.249031	1.00E-103	1	wh80g09.x1 cDNA, 3' end /clone=IMAGE:2387104
146A9	1145	1422	S63912	Hs.249247	1.00E-113	1	D10S102=FBRNP [human, fetal brain, mRNA, 3043 nt] /cds=(30,
519E8	37	628	NM_002136	Hs.249495	0	1	heterogeneous nuclear ribonucleoprotein A1
458C7	2232	2520	NM_000964	Hs.250505	1.00E-163	1	retinoic acid receptor, alpha (RARA), mRNA /cd
476A8	1060	1601	AF308285	Hs.250528	0	1	serologically defined breast cancer antigen N

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

123D7	436	2077	AL157499	Hs.250535	1.00E-153	3	mRNA; cDNA DKFZp434N2412 (from clone DKFZp434
477A10	285	370	AW291304	Hs.250600	2.00E-34	1	UI-H-BI2-agg-b-11-0-UI.s1 cDNA, 3' end /clon
172G12	726	1598	AF182420	Hs.250619	0	6	MDS019 (MDS019) mRNA, complete cds /cds=(231,1
167E11	11633	13714	NM_016252	Hs.250646	1.00E-180	2	baculoviral IAP repeat-containing 6 (BIRC6),
591E4	198	714	NM_002823	Hs.250655	4.00E-99	3	prothymosin, alpha (gene sequence 28) (PTMA),
40D9	2289	3010	M95585	Hs.250692	0	1	hepatic leukemia factor (HLF) mRNA, complete cds /cds
110D9	2336	3259	NM_003144	Hs.250773	0	3	signal sequence receptor, alpha (translocon-a
166A3	1	302	AF103458	Hs.250806	6.00E-93	2	isolate donor N clone N168K immunoglobulin kap
110C12	629	1228	M35416	Hs.250811	0	1	GTP-binding protein (RALB) mRNA, complete cds /cds=(1
458D12	1136	1714	AY007158	Hs.250820	0	1	clone CDABP0113 mRNA sequence /cds=UNKNOWN /g
177C5	658	823	J02621	Hs.251064	3.00E-32	1	non-histone chromosomal protein HMG-14 mRNA, complet
126A2	658	1009	NM_004965	Hs.251064	0	3	high-mobility group (nonhistone chromosomal)
523G1	1	337	AE000660	Hs.251465	1.00E-178	2	T-cell receptor alpha delta locus from bases 5
40G1	4	781	X72308	Hs.251526	0	3	for monocyte chemotactic protein-3 (MCP-
188G7	1	1030	NM_002789	Hs.251531	0	3	proteasome (prosome, macropain) subunit, alp
61E12	578	2275	NM_006537	Hs.251636	0	2	ubiquitin specific protease 3 (USP3), mRNA /cd
38B10	995	1211	AK026594	Hs.251653	1.00E-107	1	FLJ22941 fis, clone KAT08078, highly sim
70C3	2022	2405	X52142	Hs.251871	0	1	CTP synthetase (EC 6.3.4.2) /cds=(75,1850) /
177E9	49	406	S80990	Hs.252136	1.00E-125	2	ficolin [human, uterus, mRNA, 1736 nt] /cds=(532,1512) /gb
50F8	1841	2048	AK026712	Hs.252259	1.00E-114	15	FLJ23059 fis, clone LNG03912 /cds=(41,16
585E12	16	194	AI383340	Hs.252300	1.00E-63	1	tc76g05.x1 cDNA, 3' end /clone=IMAGE:2070584
181E12	22	99	BE963374	Hs.252338	4.00E-30	1	601657137R1 cDNA, 3' end /clone=IMAGE:3866193
477H4	290	451	AI524022	Hs.252359	8.00E-87	1	tg99f02.x1 cDNA, 3' end /clone=IMAGE:2116923
188G11	95	700	NM_007104	Hs.252574	0	2	ribosomal protein L10a (RPL10A), mRNA /cds=(1
471H9	1	285	AV706014	Hs.252580	1.00E-145	1	AV706014 cDNA, 5' end /clone=ADBAOB12 /clone_
134F9	1358	1464	AL359626	Hs.252588	5.00E-50	1	mRNA; cDNA DKFZp564F172 (from clone DKFZp564F1
597B10	13	279	NM_000981	Hs.252723	1.00E-149	28	ribosomal protein L19 (RPL19), mRNA /cds=(28,6
120D7	962	1674	NM_006054	Hs.252831	0	5	reticulon 3 (RTN3), mRNA /cds=(124,834) /gb=N
593B10	102	467	AW191929	Hs.252989	7.00E-93	1	xl77c10.x1 cDNA, 3' end /clone=IMAGE:2680722
482C11	32	122	AW195119	Hs.253151	3.00E-33	1	xn66b07.x1 cDNA, 3' end /clone=IMAGE:2699413
472C6	34	279	AW204029	Hs.253384	1.00E-137	1	UI-H-BI1-aen-d-02-0-UI.s1 cDNA, 3' end /clon
472D4	27	440	AW205624	Hs.253502	0	1	UI-H-BI1-afr-e-01-0-UI.s1 cDNA, 3' end /clon
472D1	120	362	BF750565	Hs.253550	1.00E-133	1	RC1-BN0410-261000-014-f11 cDNA /gb=BF750565
480F11	367	558	AW237483	Hs.253820	1.00E-105	1	xm72e01.x1 cDNA, 3' end /clone=IMAGE:2689752
472B5	35	363	AI432340	Hs.254006	1.00E-169	1	tg54e06.x1 cDNA, 3' end /clone=IMAGE:2112610
75E5	1	904	M14328	Hs.254105	0	5	alpha enolase mRNA, complete cds /cds=(94,1398) /gb=
592A12	1	1100	NM_001428	Hs.254105	0	5	enolase 1, (alpha) (ENO1), mRNA /cds=(94,1398)
472D10	183	414	AI364936	Hs.255100	1.00E-126	1	qz23c12.x1 cDNA, 3' end /clone=IMAGE:2027734
479H9	43	184	AW292772	Hs.255119	2.00E-70	1	UI-H-BW0-aij-d-03-0-UI.s1 cDNA, 3' end /clon
480A2	18	523	AW293267	Hs.255178	0	1	UI-H-BW0-aii-e-10-0-UI.s1 cDNA, 3' end /clon
480B7	16	298	AW293895	Hs.255249	1.00E-116	1	UI-H-BW0-ain-f-10-0-UI.s1 cDNA, 3' end /clon
479H11	23	202	AW293955	Hs.255255	3.00E-79	1	UI-H-BW0-aik-d-05-0-UI.s1 cDNA, 3' end /clon
480A4	415	598	AW294681	Hs.255336	5.00E-66	1	UI-H-BW0-ail-g-10-0-UI.s1 cDNA, 3' end /clon
480A7	223	427	AW294695	Hs.255339	1.00E-103	1	UI-H-BW0-aim-a-02-0-UI.s1 cDNA, 3' end /clon
480A8	26	338	BF514247	Hs.255340	1.00E-167	1	UI-H-BW1-ani-h-09-0-UI.s1 cDNA, 3' end /clon
480C12	239	483	AW295088	Hs.255389	1.00E-124	1	UI-H-BW0-ait-d-09-0-UI.s1 cDNA, 3' end /clon
480F9	1	423	BF531016	Hs.255390	0	1	602072345F1 cDNA, 5' end /clone=IMAGE:4215251
480B3	68	377	AW295610	Hs.255446	1.00E-161	1	UI-H-BW0-aip-c-03-0-UI.s1 cDNA, 3' end /clon

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

460H5	44	427	AA455707	Hs.255452	1.00E-161	1	aa22d09.r1 cDNA, 5' end /clone=IMAGE:814001 /
480B12	132	212	AW295664	Hs.255454	7.00E-39	1	UI-H-BW0-aip-g-12-0-UI.s1 cDNA, 3' end /clon
472E7	163	489	AI439645	Hs.255490	1.00E-166	1	tc91e08.x1 cDNA, 3' end /clone=IMAGE:2073542
480D12	84	258	AW296005	Hs.255492	8.00E-90	1	UI-H-BW0-aiu-b-01-0-UI.s1 cDNA, 3' end /clon
480F4	34	464	AW296063	Hs.255501	0	1	UI-H-BW0-aiu-g-08-0-UI.s1 cDNA, 3' end /clon
480D5	18	404	AW296490	Hs.255554	0	2	UI-H-BW0-aiq-f-08-0-UI.s1 cDNA, 3' end /clon
480E1	95	379	AW296532	Hs.255559	1.00E-101	1	UI-H-BW0-aiv-b-07-0-UI.s1 cDNA, 3' end /clon
480E5	17	326	AW296545	Hs.255560	1.00E-128	1	UI-H-BW0-aiv-c-11-0-UI.s1 cDNA, 3' end /clon
480F2	20	330	AW296730	Hs.255573	1.00E-160	1	UI-H-BW0-aix-f-12-0-UI.s1 cDNA, 3' end /clon
480G7	38	479	AW296797	Hs.255579	0	1	UI-H-BW0-ajb-e-07-0-UI.s1 cDNA, 3' end /clon
480C9	19	274	AW297339	Hs.255637	1.00E-117	1	UI-H-BW0-air-c-03-0-UI.s1 cDNA, 3' end /clon
480C4	70	191	AW297400	Hs.255647	1.00E-49	1	UI-H-BW0-ais-a-05-0-UI.s1 cDNA, 3' end /clon
480G5	17	242	AW297522	Hs.255661	2.00E-87	1	UI-H-BW0-aja-e-02-0-UI.s1 cDNA, 3' end /clon
480F10	230	560	AW294654	Hs.255687	0	1	UI-H-BW0-ail-d-10-0-UI.s1 cDNA, 3' end /clon
480G9	47	582	AW297813	Hs.255695	0	1	UI-H-BW0-aiy-g-09-0-UI.s1 cDNA, 3' end /clon
480G10	31	453	AW297827	Hs.255697	0	1	UI-H-BW0-aiy-h-11-0-UI.s1 cDNA, 3' end /clon
482G6	16	242	AW339651	Hs.255927	3.00E-78	1	he15g04.x1 cDNA, 3' end /clone=IMAGE:2919126
469B11	4	221	AW341086	Hs.256031	1.00E-99	1	xz92h04.x1 cDNA, 3' end /clone=IMAGE:2871703
140E7	2870	3589	M32315	Hs.256278	1.00E-84	2	tumor necrosis factor receptor mRNA, complete cds /cd
189H12	2839	3294	NM_001066	Hs.256278	0	2	tumor necrosis factor receptor superfamily, m
99H11	83	589	NM_005620	Hs.256290	0	4	S100 calcium-binding protein A11 (calgizzarin
58C7	1778	2264	AJ271747	Hs.256583	0	1	partial mRNA for double stranded RNA binding nu
482F4	373	628	AV719442	Hs.256959	1.00E-124	1	AV719442 cDNA, 5' end /clone=GLCBNA01 /clone_
482F5	8	377	AW440866	Hs.256961	1.00E-179	1	he05f02.x1 cDNA, 3' end /clone=IMAGE:2918139
482F8	191	315	AW440974	Hs.256971	2.00E-62	1	he06e12.x1 cDNA, 3' end /clone=IMAGE:2918254
479E7	136	567	AW444482	Hs.256979	0	2	UI-H-BI3-akb-e-05-0-UI.s1 cDNA, 3' end /clon
471H5	3	432	AI438957	Hs.257066	0	1	tc89b05.x1 cDNA, 3' end /clone=IMAGE:2073297
472G3	233	617	AW450350	Hs.257283	0	1	UI-H-BI3-akn-c-01-0-UI.s1 cDNA, 3' end /clon
472G11	112	338	AI809475	Hs.257466	1.00E-101	1	wh76d06.x1 cDNA, 3' end /clone=IMAGE:2386667
479F7	22	421	AW452467	Hs.257572	0	1	UI-H-BI3-als-e-09-0-UI.s1 cDNA, 3' end /clon
479G9	95	304	AW452513	Hs.257579	1.00E-81	1	UI-H-BW1-ame-b-03-0-UI.s1 cDNA, 3' end /clon
479F11	16	329	AW453021	Hs.257640	1.00E-163	1	UI-H-BW1-ama-c-02-0-UI.s1 cDNA, 3' end /clon
479G4	45	441	AW453044	Hs.257646	0	1	UI-H-BW1-ama-e-01-0-UI.s1 cDNA, 3' end /clon
482F9	11	256	AW467193	Hs.257667	1.00E-108	1	he07a04.x1 cDNA, 3' end /clone=IMAGE:2918286
482G2	9	271	AW467400	Hs.257680	1.00E-112	1	he10f11.x1 cDNA, 3' end /clone=IMAGE:2918637
482G8	108	428	AW467437	Hs.257682	1.00E-177	1	he17d05.x1 cDNA, 3' end /clone=IMAGE:2919273
482G12	1	417	AW467501	Hs.257687	0	1	he19e06.x1 cDNA, 3' end /clone=IMAGE:2919490
482H4	39	143	AW467746	Hs.257695	3.00E-51	1	he23d05.x1 cDNA, 3' end /clone=IMAGE:2919849
482H6	1	116	AW467863	Hs.257705	2.00E-59	1	he27c04.x1 cDNA, 3' end /clone=IMAGE:2920230
482H7	1	321	AW467864	Hs.257706	1.00E-156	1	he27c05.x1 cDNA, 3' end /clone=IMAGE:2920232
482H9	1	112	AW467992	Hs.257709	1.00E-47	1	he30b01.x1 cDNA, 3' end /clone=IMAGE:2920489
483A2	20	429	AW468207	Hs.257716	0	1	he34a12.x1 cDNA, 3' end /clone=IMAGE:2920894
483A9	11	373	AW468431	Hs.257727	0	1	he37h11.x1 cDNA, 3' end /clone=IMAGE:2921253
483B2	2	241	AW468621	Hs.257743	1.00E-119	1	he42e03.x1 cDNA, 3' end /clone=IMAGE:2921692
75B1	157	246	BE531180	Hs.258494	5.00E-44	1	601278313F1 cDNA, 5' end /clone=IMAGE:3610443
585F6	2200	4106	AL136549	Hs.258503	0	8	mRNA; cDNA DKFZp761I12121 (from clone DKFZp761
169E2	5186	5415	U20489	Hs.258609	1.00E-119	2	glomerular epithelial protein 1 (GLEPP1) comple
127A5	2142	2477	AB037790	Hs.258730	1.00E-177	1	mRNA for KIAA1369 protein, partial cds /cds=(0
171B12	4202	4314	Y10129	Hs.258742	4.00E-45	2	mybpc3 gene /cds=(33,3857) /gb=Y10129 /gi=20583
75B7	531	682	L14542	Hs.258850	3.00E-81	1	lectin-like type II integral membrane protein (NKG2-E
471G5	344	473	AI144328	Hs.259084	3.00E-61	1	oy84g04.x1 cDNA, 3' end /clone=IMAGE:1672566
479B7	73	307	AF161364	Hs.259683	1.00E-123	1	HSPC101 mRNA, partial cds /cds=(0,556) /gb=AF
146B11	1942	2174	AL136842	Hs.260024	8.00E-92	1	DKFZp434A0530 (from clone DKFZp434A

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

584A1	1085	1470	AL022398	Hs.261373	1.00E-166	1	DNA sequence from PAC 434O14 on chromosome 1q32
148B1	119	817	X60656	Hs.261802	0	2	elongation factor 1-beta /cds=(95,772)
60G3	203	3170	NM_001634	Hs.262476	0	15	S-adenosylmethionine decarboxylase 1 (AMD1)
462E7	292	374	AW300868	Hs.262789	8.00E-40	1	xk07d09.x1 cDNA, 3' end /clone=IMAGE:2666033
56F11	33	234	BF243724	Hs.263414	4.00E-82	1	601877832F1 cDNA, 5' end /clone=IMAGE:4106359
119C5	2414	2664	NM_002108	Hs.263435	1.00E-137	1	histidine ammonia-lyase (HAL), mRNA /cds=(297
105A4	3225	3775	AK025774	Hs.264190	0	3	FLJ22121 fis, clone HEP18876, highly sim
469H1	369	576	AI380111	Hs.264298	1.00E-103	1	tf98a11.x1 cDNA, 3' end /clone=IMAGE:2107292
181A3	2434	2768	NM_002535	Hs.264981	1.00E-148	2	2'-5'oligoadenylate synthetase 2 (OAS2), tra
41B7	3209	3885	M59911	Hs.265829	0	1	integrin alpha-3 chain mRNA, complete cds /cds=(73,32
75F9	264	452	AW150944	Hs.265838	2.00E-96	1	xg42e09.x1 cDNA, 3' end /clone=IMAGE:2630248
99C3	2684	3155	AK000680	Hs.266175	0	2	cDNA FLJ20673 fis, clone KAIA4464 /cds=(104,14
598E12	2417	2894	AK026669	Hs.266940	0	2	cDNA: FLJ23016 fis, clone LNG00874 /cds=UNKNOW
468B6	863	1515	NM_016569	Hs.267182	0	1	TBX3-iso protein (TBX3-iso), mRNA /cds=(116,1
115E11	1234	1713	AF271994	Hs.267288	0	1	dopamine responsive protein DRG-1 mRNA, compl
114A4	31	382	NM_024095	Hs.267400	1.00E-179	1	hypothetical protein MGC5540 (MGC5540), mRNA
166C7	1315	1919	AK001749	Hs.267604	0	2	FLJ10887 fis, clone NT2RP4002018, weakly
56A8	564	3624	AB033054	Hs.267690	0	3	for KIAA1228 protein, partial cds /cds=(0
70B10	229	2138	AK001471	Hs.268012	0	3	FLJ10609 fis, clone NT2RP2005276, highly
178D10	1831	2796	NM_012255	Hs.268555	0	2	5'-3' exoribonuclease 2 (XRN2), mRNA /cds=(68,
168B9	451	881	AF068235	Hs.268763	0	1	barrier-to-autointegration factor mRNA, com
465F2	91	433	AA613224	Hs.270264	0	1	no19d06.s1 cDNA, 3' end /clone=IMAGE:1101131
469E2	302	422	BE857296	Hs.270293	1.00E-57	1	7g27b01.x1 cDNA, 3' end /clone=IMAGE:3307657
465D10	284	405	AI270476	Hs.270341	4.00E-51	1	qu88e12.x1 cDNA, 3' end /clone=IMAGE:1979182
473F10	831	1096	AK021517	Hs.270557	1.00E-140	1	cDNA FLJ11455 fis, clone HEMBA1001497 /cds=UNK
193A10	458	563	AI818951	Hs.270614	5.00E-31	1	wj89e12.x1 cDNA, 3' end /clone=IMAGE:2410030
458E11	44	264	W03955	Hs.270717	1.00E-118	1	za62d04.r1 cDNA, 5' end /clone=IMAGE:297127 /
163C12	280	954	M30704	Hs.270833	1.00E-168	2	amphiregulin (AR) mRNA, complete cds, clones lambda-A
196F4	208	567	NM_001657	Hs.270833	1.00E-158	1	amphiregulin (schwannoma-derived growth fac
464G2	378	529	AW172850	Hs.270999	4.00E-77	1	xj04f02.x1 cDNA, 3' end /clone=IMAGE:2656251
464F5	131	476	AW572930	Hs.271264	0	1	hf17f07.x1 cDNA, 3' end /clone=IMAGE:2932165
41G6	458	880	Y16645	Hs.271387	0	1	for monocyte chemotactic protein-2 /cds=
464F2	139	220	AW975086	Hs.271420	2.00E-34	1	EST387192 cDNA /gb=AW975086 /gi=8166291 /ug=
178E10	961	1452	AK021715	Hs.271541	0	1	cDNA FLJ11653 fis, clone HEMBA1004538 /cds=UNK
129E1	73	441	NM_016049	Hs.271614	1.00E-136	1	CGI-112 protein (LOC51016), mRNA /cds=(158,78
40C9	4195	4949	X17033	Hs.271986	0	1	integrin alpha-2 subunit /cds=(48,3593) /gb
108E1	917	1331	NM_006811	Hs.272168	0	2	tumor differentially expressed 1 (TDE1), mRNA
155H10	232	715	AL021395	Hs.272279	1.00E-164	1	DNA sequence from clone RP1-269M15 on chromosome 20q1
159D3	38	238	AL034343	Hs.272295	1.00E-106	4	DNA sequence from clone RP1-108C2 on chromosome 6p12.
477C3	744	1166	AL133015	Hs.272307	0	2	mRNA; cDNA DKFZp434O2417 (from clone DKFZp434O
173D12	228	594	AL121934	Hs.272340	1.00E-140	5	DNA sequence from clone RP11-209A2 on chromosome 6. C
472D9	27	418	NM_016135	Hs.272398	0	1	transcription factor ets (TEL2), mRNA /cds=(7
465F9	1885	2345	NM_013351	Hs.272409	0	1	T-box 21 (TBX21), mRNA /cds=(211,1818) /gb=NM
41E11	1	277	NM_004167	Hs.272493	1.00E-113	1	small inducible cytokine subfamily A (Cys-Cys
462E11	8	526	NM_001503	Hs.272529	0	1	glycosylphosphatidylinositol specific phos
104C6	210	327	AE000659	Hs.272550	5.00E-61	1	T-cell receptor alpha delta locus from bases 2
596A3	411	1208	NM_013392	Hs.272736	0	5	nuclear receptor binding protein (NRBP), mRNA
75C2	1892	2188	AK000316	Hs.272793	1.00E-165	1	FLJ20309 fis, clone HEP07296 /cds=(41,127
58C6	1	956	NM_006009	Hs.272897	0	2	Tubulin, alpha, brain-specific (TUBA3), mRNA

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

190H8	3246	3771	AK024471	Hs.273230	1.00E-165	2	mRNA for FLJ00064 protein, partial cds /cds=(0
590E11	1512	1860	NM_014230	Hs.273307	1.00E-168	4	signal recognition particle 68kD (SRP68), mRN
588H2	696	1454	NM_000516	Hs.273385	0	3	guanine nucleotide binding protein (G protein)
165E9	3186	3695	NM_014871	Hs.273397	0	1	KIAA0710 gene product (KIAA0710), mRNA /cds=(
462A6	394	496	AA527312	Hs.273775	2.00E-42	1	ng36a08.s1 cDNA, 3' end /clone=IMAGE:936854 /
587F1	1763	1978	AL050353	Hs.274170	1.00E-112	1	mRNA; cDNA DKFZp564C0482 (from clone DKFZp564C
177E5	1448	1876	AK000765	Hs.274248	0	1	FLJ20758 fis, clone HEP01508 /cds=(464,13
59E7	1	301	AF151049	Hs.274344	1.00E-159	3	HSPC215 mRNA, complete cds /cds=(92,451) /gb=
174A6	931	1352	NM_004301	Hs.274350	0	1	BAF53 (BAF53A), mRNA /cds=(136,1425) /gb=NM_0
99E2	718	1391	NM_018477	Hs.274369	0	4	uncharacterized hypothalamus protein HARP11
117F6	3046	3478	AB037844	Hs.274396	0	2	mRNA for KIAA1423 protein, partial cds /cds=(0
52F3	1724	2342	NM_005346	Hs.274402	1.00E-149	48	heat shock 70kD protein 1 (HSPA1B), mRNA /cds=(
516B1	719	1026	NM_018975	Hs.274428	1.00E-161	2	TRF2-interacting telomeric RAP1 protein (RAP
104A1	1943	2396	AK002127	Hs.274439	0	1	FLJ11265 fis, clone PLACE1009158 /cds=(30
137D6	1697	1817	NM_001403	Hs.274466	8.00E-49	1	eukaryotic translation elongation factor 1 a
108D11	321	646	X16863	Hs.274467	1.00E-160	1	Fc-gamma RIII-1 cDNA for Fc-gamma receptor III-1 (CD
107F1	567	895	AF283771	Hs.274472	1.00E-168	1	clone TCBAP0774 mRNA sequence /cds=UNKNOWN /g
517B9	4	480	NM_002128	Hs.274472	0	3	high-mobility group (nonhistone chromosomal)
514C8	254	539	M12888	Hs.274474	1.00E-144	2	T-cell receptor germline beta-chain gene C-region C-
460G5	602	775	M12679	Hs.274485	3.00E-94	1	Cw1 antigen mRNA, complete cds /cds=(0,617) /gb=M1267
463G7	163	744	D90145	Hs.274535	0	4	LD78 beta gene /cds=(86,367) /gb=D90145 /gi=219907 /
472E10	277	391	AI393960	Hs.274851	6.00E-59	1	tg11d04.x1 cDNA, 3' end /clone=IMAGE:2108455
115A11	156	446	NM_014624	Hs.275243	1.00E-157	8	S100 calcium-binding protein A6 (calcyclin) (
102C6	23	448	AA610514	Hs.275611	1.00E-161	1	np93h02.s1 /clone=IMAGE:1133907 /gb=AA6
160E3	24	304	AA757952	Hs.275773	1.00E-74	3	zg49e07.s1 3' end /clone=IMAGE:396708 /
500B8	26	536	NM_022551	Hs.275865	0	3	ribosomal protein S18 (RPS18), mRNA /cds=(46,5
522D9	184	593	NM_001959	Hs.275959	0	1	eukaryotic translation elongation factor 1 b
151H4	1	196	AA984890	Hs.276063	5.00E-58	1	am62e06.s1 cDNA, 3' end /clone=IMAGE:1576642
476B10	362	615	BF510670	Hs.276341	1.00E-116	1	UI-H-BI4-aof-b-08-0-UI.s1 cDNA, 3' end /clon
144F10	73	279	AI318342	Hs.276662	8.00E-57	1	ta73c09.x1 3' end /clone=IMAGE:2049712
593G1	17	88	BE747210	Hs.276718	2.00E-26	1	601580926F1 cDNA, 5' end /clone=IMAGE:3929430
473E3	205	488	AI380791	Hs.276766	1.00E-144	1	tg04b12.x1 cDNA, 3' end /clone=IMAGE:2107775
598A2	72	427	NM_001803	Hs.276770	0	19	CDW52 antigen (CAMPATH-1 antigen) (CDW52), mR
170H2	83	432	X62466	Hs.276770	0	1	CAMPATH-1 (CDw52) antigen /cds=(33,218)
464F7	2	454	AI492640	Hs.276903	0	2	qz18a06.x1 cDNA, 3' end /clone=IMAGE:2021842
464E5	102	191	AI493726	Hs.276907	3.00E-44	2	qz12f08.x1 cDNA, 3' end /clone=IMAGE:2021319
50B5	42	308	AI581383	Hs.276988	5.00E-77	1	to71c02.x1 cDNA, 3' end /clone=IMAGE:2183714
468C6	40	279	AI740667	Hs.277201	1.00E-64	1	wg07b07.x1 cDNA, 3' end /clone=IMAGE:2364373
111D12	1	562	AI749435	Hs.277224	1.00E-118	9	at24b04.x1 cDNA, 3' end /clone=IMAGE:2356015
459B4	176	367	AI811065	Hs.277293	2.00E-38	1	tr03f05.x1 cDNA, 3' end /clone=IMAGE:2217249
477H3	6227	6584	NM_013449	Hs.277401	1.00E-132	1	bromodomain adjacent to zinc finger domain, 2A
54A8	34	301	AW050975	Hs.277672	3.00E-48	1	wz25f04.x1 cDNA, 3' end /clone=IMAGE:2559103
459E4	1532	2061	NM_006389	Hs.277704	0	1	oxygen regulated protein (150kD) (ORP150), mR
109B6	3281	3721	U65785	Hs.277704	0	1	150 kDa oxygen-regulated protein ORP150 mRNA, complet
524H7	2979	3350	NM_005899	Hs.277721	0	1	membrane component, chromosome 17, surface ma
472F10	425	556	AW082714	Hs.277738	5.00E-69	1	xb61f07.x1 cDNA, 3' end /clone=IMAGE:2580805
176D1	113	269	AW262728	Hs.277994	6.00E-32	1	xq94a12.x1 cDNA, 3' end /clone=IMAGE:2758270
464H4	2138	3563	NM_016733	Hs.278027	0	9	LIM domain kinase 2 (LIMK2), transcript varian

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

145C9	533	1446	D13316	Hs.278238	0	3	transcription factor, E4TF1-47, complete cds
161C3	339	560	NM_002041	Hs.278238	1.00E-123	1	GA-binding protein transcription factor, bet
74C9	345	1048	AK026632	Hs.278242	0	3	FLJ22979 fis, clone KAT11379, highly sim
59E2	255	782	L24804	Hs.278270	0	2	(p23) mRNA, complete cds /cds=(232,714) /gb=L24804 /
521H10	8	461	AI720536	Hs.278302	1.00E-114	4	as83c02.x1 cDNA, 3' end /clone=IMAGE:2335298
118C6	830	1104	NM_001995	Hs.278333	1.00E-148	1	fatty-acid-Coenzyme A ligase, long-chain 1 (
104E9	248	417	AF151054	Hs.278429	2.00E-78	1	HSPC220 mRNA, complete cds /cds=(288,818) /gb
594F10	379	1760	NM_016520	Hs.278429	0	4	hepatocellular carcinoma-associated antigen
126D11	7374	7716	NM_006289	Hs.278559	0	1	talin (TLN), mRNA /cds=(126,7751) /gb=NM_0062
589E6	3078	5778	NM_003105	Hs.278571	0	3	sortilin-related receptor, L(DLR class) A re
102C10	669	1180	D14041	Hs.278573	0	1	for H-2K binding factor-2, complete cds /
526H8	167	4709	NM_015874	Hs.278573	0	5	H-2K binding factor-2 (LOC51580), mRNA /cds=(
120A12	732	1305	AB029031	Hs.278586	0	1	mRNA for KIAA1108 protein, partial cds /cds=(0
126F4	3138	3515	AF035737	Hs.278589	0	2	general transcription factor 2-I (GTF2I) mRNA
40A7	3179	3864	U24578	Hs.278625	0	1	RP1 and complement C4B precursor (C4B) genes, partial
50C4	4401	4581	AB002334	Hs.278671	2.00E-60	1	KIAA0336 gene, complete cds /cds=(253,5004)
106E12	104	1222	D50525	Hs.278693	0	11	TI-227H /cds=UNKNOWN /gb=D50525 /gi=1167502
467E10	168	542	BE973840	Hs.278704	1.00E-145	1	601680647F1 cDNA, 5' end /clone=IMAGE:3951154
75F2	1121	1772	J04755	Hs.278718	0	37	ferritin H processed pseudogene, complete cds /cds=UN
170E12	204	843	AL121735	Hs.278736	0	2	Isoform of human GTP-binding protein G25K /cds=(104,679) /
103F4	589	926	NM_019597	Hs.278857	0	1	heterogeneous nuclear ribonucleoprotein H2
37F8	3	519	U01923	Hs.278857	0	1	BTK region clone ftp-3 mRNA /cds=UNKNOWN /gb=U01923 /
66B11	2195	2512	AB029027	Hs.279039	1.00E-172	1	for KIAA1104 protein, complete cds /cds=(
171G3	219	815	AK027258	Hs.279040	0	2	FLJ23605 fis, clone LNG15982, highly sim
172E12	18	95	NM_014065	Hs.279040	4.00E-27	2	HT001 protein (HT001), mRNA /cds=(241,1203) /
596A12	1	225	BE220869	Hs.279231	2.00E-78	1	hu01g02.x1 cDNA, 3' end /clone=IMAGE:3165362
61H2	20	220	BE279328	Hs.279429	2.00E-32	3	601157666F1 cDNA, 5' end /clone=IMAGE:3504328
458E12	1835	2473	NM_014160	Hs.279474	0	1	HSPC070 protein (HSPC070), mRNA /cds=(331,158
110F3	983	1614	NM_016160	Hs.279518	0	1	amyloid precursor protein homolog HSD-2 (LOC5
37E5	39	732	AK001403	Hs.279521	0	1	FLJ10541 fis, clone NT2RP2001381 /cds=(3
66D6	6	463	BE502919	Hs.279522	0	1	hz81b08.x1 cDNA, 3' end /clone=IMAGE:3214359
123A11	411	903	NM_013237	Hs.279529	0	2	px19-like protein (PX19), mRNA /cds=(176,835)
185A10	809	1324	NM_002817	Hs.279554	0	1	proteasome (prosome, macropain) 26S subunit,
472H9	88	543	AL582047	Hs.279555	0	1	AL582047 cDNA /clone=CS0DL003YD01-(3-prime)
41A2	1	326	AK000575	Hs.279581	1.00E-162	1	FLJ20568 fis, clone REC00775 /cds=(6,422)
135F4	648	935	NM_016283	Hs.279586	1.00E-110	1	adrenal gland protein AD-004 (LOC51578), mRNA
69D9	841	935	D16217	Hs.279607	9.00E-40	1	calpastatin, complete cds /cds=(162,2288) /
116B6	938	1562	NM_001750	Hs.279607	0	1	calpastatin (CAST), mRNA /cds=(66,1358) /gb=
473F4	6847	7401	NM_007329	Hs.279611	0	1	deleted in malignant brain tumors 1 (DMBT1), tr
123C7	2488	2684	NM_021644	Hs.279681	1.00E-105	1	heterogeneous nuclear ribonucleoprotein H3
586E2	357	633	NM_014169	Hs.279761	3.00E-97	1	HSPC134 protein (HSPC134), mRNA /cds=(45,716)
464D6	383	524	NM_016154	Hs.279771	1.00E-33	1	ras-related GTP-binding protein 4b (RAB4B), m
99G9	1375	1835	NM_013388	Hs.279784	0	1	prolactin regulatory element binding (PREB),
590F4	1045	1540	NM_003883	Hs.279789	0	2	histone deacetylase 3 (HDAC3), mRNA /cds=(55,1
163E1	59	564	NM_015932	Hs.279813	0	3	hypothetical protein (HSPC014), mRNA /cds=(8
525G5	3914	4160	NM_014819	Hs.279849	1.00E-138	1	KIAA0438 gene product (KIAA0438), mRNA /cds=(
598A10	9	821	NM_003295	Hs.279860	0	19	tumor protein, translationally-controlled 1
526C8	734	1166	NM_016007	Hs.279867	0	1	CGI-59 protein (LOC51625), mRNA /cds=(2,1153)
183G12	758	1093	NM_017774	Hs.279893	0	1	hypothetical protein FLJ20342 (FLJ20342), mR
36B3	247	611	AK025623	Hs.279901	0	1	FLJ21970 fis, clone HEP05733, highly sim
592G3	479	1052	NM_016146	Hs.279901	0	4	PTD009 protein (PTD009), mRNA /cds=(257,916)

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

38F5	811	1256	AF151875	Hs.279918	0	4	CGI-117 protein mRNA, complete cds /cds=(456,9
161E3	542	862	NM_016391	Hs.279918	1.00E-151	1	hypothetical protein (HSPC111), mRNA /cds=(6
584F11	10	212	NM_014248	Hs.279919	1.00E-112	2	ring-box 1 (RBX1), mRNA /cds=(6,332) /gb=NM_0
588H7	400	1155	NM_003404	Hs.279920	0	12	tyrosine 3-monooxygenase/tryptophan 5-monoo
169C8	400	1155	X57346	Hs.279920	1.00E-131	2	HS1 protein /cds=(372,1112) /gb=X57346
147A1	209	1978	AK025927	Hs.279921	0	8	FLJ22274 fis, clone HRC03616, highly sim
591H11	48	1810	NM_016127	Hs.279921	1.00E-176	33	HSPC035 protein (LOC51669), mRNA /cds=(16,103
69D1	727	1776	NM_014366	Hs.279923	0	3	putative nucleotide binding protein, estradio
52C6	303	1151	V00522	Hs.279930	0	2	encoding major histocompatibility complex gene
158C11	2483	2785	D84224	Hs.279946	1.00E-166	2	methionyl tRNA synthetase, complete c
194E7	1525	1767	NM_004990	Hs.279946	1.00E-125	1	methionine-tRNA synthetase (MARS), mRNA /cds
62E5	215	701	U93243	Hs.279948	0	1	Ubc6p homolog mRNA, complete cds /cds=(27,983)
145G3	1	1882	AK024090	Hs.281434	1.00E-147	5	FLJ14028 fis, clone HEMBA1003838 /cds=UN
473A6	1	310	BE552131	Hs.282091	1.00E-158	1	hw29b05.x1 cDNA, 3' end /clone=IMAGE:3184305
52C12	1	455	R67739	Hs.282401	0	1	yi28c06.r1 cDNA, 5' end /clone=IMAGE:140554 /
112A3	5072	5274	NM_006165	Hs.282441	3.00E-83	1	nuclear factor related to kappa B binding prote
61H3	443	577	AV648638	Hs.282867	2.00E-68	4	AV648638 cDNA, 3' end /clone=GLCBLE12 /clone_
37D3	38	766	AF287008	Hs.283022	0	5	triggering receptor expressed on monocytes 1
125C5	32	748	NM_018643	Hs.283022	0	13	triggering receptor expressed on myeloid cell
41B1	597	1084	NM_018636	Hs.283106	0	2	hypothetical protein PRO2987 (PRO2987), mRNA
111E9	1111	1405	AB037802	Hs.283109	1.00E-152	1	mRNA for KIAA1381 protein, partial cds /cds=(0
169D7	5	175	BE672733	Hs.283216	2.00E-37	1	7b75g07.x1 3' end /clone=IMAGE:3234108
74G11	47	384	BE676472	Hs.283267	1.00E-151	1	7f30c05.x1 cDNA, 3' end /clone=IMAGE:3296168
191A5	256	890	NM_018507	Hs.283330	0	3	hypothetical protein PRO1843 (PRO1843), mRNA
465B7	114	638	AW979262	Hs.283410	0	2	EST391372 cDNA /gb=AW979262 /gi=8170550 /ug=
143E1	1970	2258	NM_020217	Hs.283611	1.00E-110	1	hypothetical protein DKFZp547I014 (DKFZp547I
54E9	385	739	AF116620	Hs.283630	0	3	PRO1068 mRNA, complete cds /cds=UNKNOWN /gb=A
462D10	63	279	NM_007220	Hs.283646	1.00E-119	1	carbonic anhydrase VB, mitochondrial (CA5B),
518B11	359	690	NM_016056	Hs.283670	1.00E-167	2	CGI-119 protein (LOC51643), mRNA /cds=(0,776)
36H5	1	226	BE778549	Hs.283674	8.00E-85	1	601466063F1 cDNA, 5' end /clone=IMAGE:3869391
126H10	907	1431	NM_017801	Hs.283685	0	1	hypothetical protein FLJ20396 (FLJ20396), mR
69B1	2288	3232	AF103803	Hs.283690	0	6	clone H41 unknown mRNA /cds=(323,1099) /gb=AF
98B1	162	489	NM_018476	Hs.283719	1.00E-110	1	uncharacterized hypothalamus protein HBEX2
39C3	997	3088	NM_020151	Hs.283722	0	2	GTT1 protein (GTT1), mRNA /cds=(553,1440) /gb
592E4	13	2219	NM_020357	Hs.283728	0	2	PEST-containing nuclear protein (pcnp), mRNA
142F11	138	371	AF173296	Hs.283740	1.00E-130	1	e(y)2 homolog mRNA, complete cds /cds=(216,521
592F3	480	858	NM_013234	Hs.283781	0	2	muscle specific gene (M9), mRNA /cds=(171,827)
159E5	3	281	AL121916	Hs.283838	1.00E-113	6	DNA sequence from clone RP1-189G13 on chromosome 20.
142H10	517	892	AL121585	Hs.283864	9.00E-70	2	DNA sequence from clone RP11-504H3 on chromosome 20 C
166D3	1	227	X72475	Hs.283972	6.00E-70	1	for rearranged Ig kappa light chain variable
134E8	980	1302	NM_014110	Hs.284136	0	47	PRO2047 protein (PRO2047), mRNA /cds=(798,968
596C5	30	705	NM_006134	Hs.284142	0	2	chromosome 21 open reading frame 4 (C21ORF4), m
74A4	1944	2157	AL359585	Hs.284158	1.00E-110	3	cDNA DKFZp762B195 (from clone DKFZp762B1
159A4	159	1414	AF165521	Hs.284162	0	4	ribosomal protein L30 isolog (L30) mRNA, compl
597F9	836	1000	NM_016304	Hs.284162	1.00E-88	1	60S ribosomal protein L30 isolog (LOC51187), m
462D2	655	1306	NM_016301	Hs.284164	0	1	protein x 0004 (LOC51184), mRNA /cds=(31,885)
458C6	720	910	AP001753	Hs.284189	1.00E-102	1	genomic DNA, chromosome 21q, section 97/105 /
165D5	1482	2302	AB040120	Hs.284205	0	2	mRNA for BCG induced integral membrane protein

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

180C12	309	602	BF381953	Hs.284235	1.00E-148	2	601816251F1 cDNA, 5' end /clone=IMAGE:4050061
67D9	27	2026	AK024969	Hs.284249	0	10	FLJ21316 fis, clone COL02253, highly sim
39D1	307	2899	U90552	Hs.284283	0	5	butyrophilin (BTF5) mRNA, complete cds /cds=(359,190
147C8	391	556	AF161451	Hs.284295	2.00E-58	1	HSPC333 mRNA, partial cds /cds=(0,443) /gb=AF
192C12	333	484	AV700210	Hs.284605	5.00E-57	1	AV700210 cDNA, 3' end /clone=GKBALC03 /clone_
49G11	380	523	AV700636	Hs.284674	4.00E-33	1	AV700636 cDNA, 3' end /clone=GKBAGH12 /clone_
115C11	375	1001	AK023291	Hs.285017	0	1	cDNA FLJ13229 fis, clone OVARC1000106 /cds=(15
458H8	1544	2233	AK023459	Hs.285107	0	1	cDNA FLJ13397 fis, clone PLACE1001351 /cds=(22
70F4	11	605	AV700298	Hs.285173	0	4	AV700298 cDNA, 3' end /clone=GKCBVGO5 /clone_
66C6	684	1415	NM_001300	Hs.285313	0	5	core promoter element binding protein (COPEB),
169F2	4	460	BF684382	Hs.285555	0	2	602141836F1 5' end /clone=IMAGE:4302776
171F12	646	839	X58529	Hs.285823	6.00E-99	2	rearranged immunoglobulin mRNA for mu heavy chain enh
142F10	1438	1728	AK025788	Hs.285833	1.00E-152	1	FLJ22135 fis, clone HEP20858 /cds=UNKNOW
171H2	1	2500	AL050376	Hs.285853	5.00E-21	1	mRNA; cDNA DKFZp586J101 (from clone DKFZp586J1
40C5	786	1163	AK026603	Hs.286124	0	2	FLJ22950 fis, clone KAT09618, highly sim
458D9	55	684	NM_016041	Hs.286131	0	1	CGI-101 protein (LOC51009), mRNA /cds=(6,635)
458D1	1	310	AK025886	Hs.286194	1.00E-151	1	cDNA: FLJ22233 fis, clone HRC02016 /cds=(35,12
515C10	817	1136	AK021791	Hs.286212	1.00E-138	1	cDNA FLJ11729 fis, clone HEMBA1005394, modera
71C7	285	2441	AK026933	Hs.286236	0	7	cDNA: FLJ23280 fis, clone HEP07194 /cds=(468,1
184B9	372	612	BE965319	Hs.286754	3.00E-66	2	601659229R1 cDNA, 3' end /clone=IMAGE:3895783
586C12	18	381	NM_000996	Hs.287361	0	3	ribosomal protein L35a (RPL35A), mRNA /cds=(6
36C6	152	685	AJ277247	Hs.287369	0	37	for interleukin 21 (IL-21 gene) /cds=(71,
513H8	17	690	NM_020525	Hs.287369	0	510	interleukin 22 (IL22), mRNA /cds=(71,610) /gb
586G2	3978	4107	NM_021621	Hs.287387	3.00E-68	1	caspase recruitment domain protein 7 (CARD7),
99D12	2330	2851	NM_015906	Hs.287414	0	1	transcriptional intermediary factor 1 gamma (
182A2	284	576	AK024331	Hs.287631	1.00E-156	1	cDNA FLJ14269 fis, clone PLACE1003864 /cds=UN
465A11	2226	2321	AK024372	Hs.287634	1.00E-42	1	cDNA FLJ14310 fis, clone PLACE3000271 /cds=(40
190A11	679	1126	AK026769	Hs.287725	0	1	cDNA: FLJ23116 fis, clone LNG07945, highly sim
75E2	479	837	AL390738	Hs.287788	1.00E-146	3	DNA sequence from clone RP11-438F9 on chromosome 13 C
59B7	488	1071	AK022537	Hs.287863	0	1	FLJ12475 fis, clone NT2RM1000962 /cds=(16
460E8	1611	1979	AK024092	Hs.287864	0	1	cDNA FLJ14030 fis, clone HEMBA1004086 /cds=UNK
465F11	5714	6271	NM_006312	Hs.287994	0	1	nuclear receptor co-repressor 2 (NCOR2), mRNA
150E12	2041	2720	AK026834	Hs.287995	0	3	FLJ23181 fis, clone LNG11094 /cds=UNKNOW
52D9	703	1482	AB016247	Hs.288031	0	1	for sterol-C5-desaturase, complete cds
37F4	1091	1655	AK025375	Hs.288061	1.00E-141	20	FLJ21722 fis, clone COLF0522, highly sim
188G5	1081	1753	NM_001101	Hs.288061	0	69	actin, beta (ACTB), mRNA /cds=(73,1200) /gb=N
171C12	2103	2426	AB046857	Hs.288140	1.00E-158	1	KIAA1637 protein, partial cds /cds=(0
104E8	1354	1790	AK023078	Hs.288141	0	1	FLJ13016 fis, clone NT2RP3000624, modera
181A4	1890	2507	AK022030	Hs.288178	0	2	cDNA FLJ11968 fis, clone HEMBB1001133 /cds=UNK
129A1	3522	3748	J04144	Hs.288204	1.00E-125	1	angiotensin I-converting enzyme mRNA, complete cds /
598D12	1464	1947	AK025643	Hs.288224	0	3	cDNA: FLJ21990 fis, clone HEP06386 /cds=(22,49
52E6	920	1388	AK023402	Hs.288416	0	2	FLJ13340 fis, clone OVARC1001942, weakly
165E3	303	640	NM_020666	Hs.288417	0	1	protein serine threonine kinase Clk4 (CLK4),
53D3	1	153	AK022280	Hs.288435	6.00E-76	1	FLJ12218 fis, clone MAMMA1001075, modera
586C2	223	448	BF110312	Hs.288443	1.00E-63	3	7n36d08.x1 cDNA, 3' end /clone=IMAGE:3566654

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

521F12	1922	2248	AK026923	Hs.288455	0	1	cDNA: FLJ23270 fis, clone COL10309, highly sim
120A11	825	1855	AK026078	Hs.288555	0	2	cDNA: FLJ22425 fis, clone HRC08686 /cds=UNKNOW
129D11	1723	1984	AK023470	Hs.288673	1.00E-143	2	FLJ13408 fis, clone PLACE1001672, weakly
109B12	1686	2086	AK025215	Hs.288708	1.00E-121	8	FLJ21562 fis, clone COL06420 /cds=(238,2
178F11	387	558	NM_005402	Hs.288757	3.00E-93	1	v-ral simian leukemia viral oncogene homolog
58F8	1262	1604	AK022735	Hs.288836	0	1	cDNA FLJ12673 fis, clone NT2RM4002344 /cds=(2,
163E11	360	1687	AK024094	Hs.288856	1.00E-25	2	FLJ14032 fis, clone HEMBA1004353, highly
105B4	741	1243	AK025092	Hs.288872	0	1	FLJ21439 fis, clone COL04352 /cds=(206,1
106D10	1598	2291	AB014515	Hs.288891	0	3	for KIAA0615 protein, complete cds /cds=(
460F8	154	2487	NM_021818	Hs.288906	1.00E-150	2	WW Domain-Containing Gene (WW45), mRNA /cds=(
48A6	560	1258	NM_017644	Hs.288922	0	1	hypothetical protein FLJ20059 (FLJ20059), mR
168B10	1271	1747	AK023320	Hs.288929	0	1	FLJ13258 fis, clone OVARC1000862, modera
114E2	2395	2849	AK023256	Hs.288932	0	1	cDNA FLJ13194 fis, clone NT2RP3004378, weakly
586F9	368	730	AK026363	Hs.288936	1.00E-162	4	cDNA: FLJ22710 fis, clone HSI13340 /cds=UNKNOW
180B4	831	959	NM_000344	Hs.288986	1.00E-32	1	survival of motor neuron 1, telomeric (SMN1),
149A12	10	1958	AK025467	Hs.289008	0	5	FLJ21814 fis, clone HEP01068 /cds=UNKNOW
117B5	5160	5611	NM_012231	Hs.289024	1.00E-141	1	PR domain containing 2, with ZNF domain (PRDM2)
469A5	3132	3365	AK024456	Hs.289034	1.00E-106	1	mRNA for FLJ00048 protein, partial cds /cds=(2
461F6	396	473	AK024197	Hs.289037	7.00E-37	1	cDNA FLJ14135 fis, clone MAMMA1002728 /cds=UN
176G11	1049	1811	AK024669	Hs.289069	0	4	cDNA: FLJ21016 fis, clone CAE05735 /cds=(90,11
473A5	1343	1937	NM_013326	Hs.289080	0	1	colon cancer-associated protein Mic1 (MIC1),
591G2	14	2259	NM_005348	Hs.289088	0	14	heat shock 90kD protein 1, alpha (HSPCA), mRNA
70D3	21	2912	X15183	Hs.289088	0	17	90-kDa heat-shock protein /cds=(60,2258) /g
37E8	780	1509	AK026033	Hs.289092	0	5	FLJ22380 fis, clone HRC07453, highly sim
74B10	408	791	X00453	Hs.289095	1.00E-153	2	gene fragment for DX alpha-chain signal peptide,
518B5	870	1128	NM_005313	Hs.289101	1.00E-119	1	glucose regulated protein, 58kD (GRP58), mRNA
472A3	116	304	X83300	Hs.289103	4.00E-84	1	H.sapiens SMA4 mRNA /cds=(66,488) /gb=X83300 /gi=603028 /
112G6	1703	2550	NM_001166	Hs.289107	0	5	baculoviral IAP repeat-containing 2 (BIRC2),
37F11	1996	2580	U37547	Hs.289107	0	2	IAP homolog B (MIHB) mRNA, complete cds /cds=(1159,301
169A12	371	588	X57812	Hs.289110	2.00E-84	1	rearranged immunoglobulin lambda light chain /c
472D6	2102	2424	AF294900	Hs.289118	1.00E-121	1	beta, beta-carotene 15,15'- dioxygenase (BCD
151D1	2214	2294	AK025846	Hs.289721	1.00E-38	2	FLJ22193 fis, clone HRC01108 /cds=UNKNOW
40A8	160	346	AI761924	Hs.289834	2.00E-94	1	wg68h03.x1 cDNA, 3' end /clone=IMAGE:2370293
468D5	42	105	AA719103	Hs.290535	5.00E-29	1	zh33d10.s1 cDNA, 3' end /clone=IMAGE:413875 /
515B6	7	249	AA837754	Hs.291129	2.00E-61	1	oe10d02.s1 cDNA /clone=IMAGE:1385475 /gb=AA
594C9	16	319	NM_005745	Hs.291904	1.00E-150	1	accessory proteins BAP31/BAP29 (DXS1357E), m
476C10	180	311	AI184710	Hs.292276	8.00E-62	1	qd64a01.x1 cDNA, 3' end /clone=IMAGE:1734216
466G5	65	431	AA461604	Hs.292451	0	1	zx51d08.r1 cDNA, 5' end /clone=IMAGE:795759 /
331F12	142	314	BF310166	Hs.292457	3.00E-85	1	601894826F1 cDNA, 5' end /clone=IMAGE:4124119
590D6	1	406	BG339050	Hs.292457	0	2	602436875F1 cDNA, 5' end /clone=IMAGE:4554643
150G5	160	431	AI440234	Hs.292490	6.00E-66	1	ti99h12.x1 cDNA, 3' end /clone=IMAGE:2140199
594F8	319	447	AA761571	Hs.292519	1.00E-57	1	nz23d06.s1 cDNA, 3' end /clone=IMAGE:1288619
122E2	91	307	AI582954	Hs.292553	4.00E-47	1	tr98e07.x1 cDNA, 3' end /clone=IMAGE:2227140
41E5	363	463	D59502	Hs.292590	3.00E-48	1	HUM041H11A cDNA, 3' end /clone=GEN-041H11 /cl
99B8	215	378	AI672433	Hs.292615	6.00E-62	4	wa03b05.x1 cDNA, 3' end /clone=IMAGE:2296977
72C6	198	484	AA719537	Hs.292877	1.00E-112	3	zh40g12.s1 cDNA, 3' end /clone=IMAGE:414598 /
157H5	49	447	AI962127	Hs.292901	1.00E-126	1	wx77f07.x1 3' end /clone=IMAGE:2549701
115C2	2052	2613	NM_006310	Hs.293007	0	1	aminopeptidase puromycin sensitive (NPEPPS),
463F3	14	445	AW629485	Hs.293352	0	2	hi59b07.x1 cDNA, 3' end /clone=IMAGE:2976565
193H8	94	333	AI263141	Hs.293444	7.00E-58	1	qw90c01.x1 cDNA, 3' end /clone=IMAGE:1998336
170G9	46	713	AI452611	Hs.293473	9.00E-21	1	tj27g07.x1 cDNA, 3' end /clone=IMAGE:2142780

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

100F9	554	666	BE905040	Hs.293515	2.00E-43	1	601496859F1 cDNA, 5' end /clone=IMAGE:3898767
588G9	153	507	BF794089	Hs.293658	1.00E-143	1	602255649F1 cDNA, 5' end /clone=IMAGE:4338732
142G8	2	231	AV701332	Hs.293689	1.00E-79	1	AV701332 cDNA, 5' end /clone=ADAABD03 /clone_
137A4	1	557	BF029654	Hs.293777	0	1	601765621F1 cDNA, 5' end /clone=IMAGE:3997900
478C6	442	622	BE748123	Hs.293842	3.00E-63	1	601571679F1 cDNA, 5' end /clone=IMAGE:3838675
100E7	198	488	BE748663	Hs.293842	1.00E-145	1	601571679T1 cDNA, 3' end /clone=IMAGE:3838675
110B4	246	469	NM_016398	Hs.293905	1.00E-122	1	hypothetical protein (HSPC131), mRNA /cds=(1
466D2	198	543	AW972477	Hs.294083	1.00E-180	1	EST384568 cDNA /gb=AW972477 /gi=8162323 /ug=
100C10	1	398	AW963235	Hs.294092	0	2	EST375308 /gb=AW963235 /gi=8153071 /ug=
118F10	418	552	BF245076	Hs.294110	1.00E-48	1	601863910F1 cDNA, 5' end /clone=IMAGE:4082235
596H2	1150	2308	BC002450	Hs.294135	0	20	ribosomal protein L4, clone MGC:776, mRNA, co
596B4	139	414	BE621121	Hs.294309	7.00E-73	3	601493943F1 cDNA, 5' end /clone=IMAGE:3896051
114D4	600	738	BE961923	Hs.294348	8.00E-33	1	601655335R1 cDNA, 3' end /clone=IMAGE:3845768
66D11	185	625	BE963811	Hs.294578	1.00E-127	6	601657462R1 cDNA, 3' end /clone=IMAGE:3875846
53E11	433	701	BE964149	Hs.294612	5.00E-81	1	601657833R1 cDNA, 3' end /clone=IMAGE:3875984
179A11	442	776	BF313856	Hs.294754	9.00E-79	1	601902261F1 5' end /clone=IMAGE:4134998
102B9	146	347	H71236	Hs.295055	7.00E-90	2	ys12f10.s1 cDNA, 3' end /clone=IMAGE:214603 /
110F4	136	358	H80108	Hs.295107	1.00E-118	1	yu09f02.s1 cDNA, 3' end /clone=IMAGE:233307 /
593F2	78	381	AF212224	Hs.295231	1.00E-172	3	CLK4 mRNA, complete cds /cds=(153,1514) /gb=A
50G9	355	415	AI052431	Hs.295451	1.00E-26	2	oz07e08.x1 cDNA, 3' end /clone=IMAGE:1674662
102E4	99	413	AI560651	Hs.295682	1.00E-146	8	tq60f01.x1 cDNA, 3' end /clone=IMAGE:2213209
486F7	263	489	BF572855	Hs.295806	1.00E-100	1	602079424F2 cDNA, 5' end /clone=IMAGE:4254172
39C1	2054	2315	AL050141	Hs.295833	1.00E-144	6	cDNA DKFZp586O031 (from clone DKFZp586O0
192D3	48	551	AW081320	Hs.295945	1.00E-158	4	xc30f12.x1 cDNA, 3' end /clone=IMAGE:2585807
102B7	753	850	AL117536	Hs.295969	5.00E-39	1	cDNA DKFZp434G012 (from clone DKFZp434G0
168D1	73	1193	AL360190	Hs.295978	1.00E-134	3	mRNA full length insert cDNA clone EUROIMAGE 74
47D6	103	331	AW150085	Hs.295997	3.00E-79	8	xg36f04.x1 cDNA, 3' end /clone=IMAGE:2629663
151H9	197	507	AW264291	Hs.296057	1.00E-113	1	xq97g08.x1 cDNA, 3' end /clone=IMAGE:2758622
56A1	1034	1220	AJ012504	Hs.296151	3.00E-74	1	activated in tumor suppression, clone TSA
525D12	42	545	AI922889	Hs.296159	1.00E-148	42	wn64g11.x1 cDNA, 3' end /clone=IMAGE:2450276
72C12	280	545	AW166001	Hs.296159	1.00E-84	10	xf43e11.x1 cDNA, 3' end /clone=IMAGE:2620844
99B1	21	286	BE259480	Hs.296183	4.00E-81	3	601106571F1 cDNA, 5' end /clone=IMAGE:3342929
143F5	18	178	BE962588	Hs.296183	1.00E-55	1	601655929R1 cDNA, 3' end /clone=IMAGE:3855823
110A10	2115	2237	AL096752	Hs.296243	1.00E-61	1	cDNA DKFZp434A012 (from clone DKFZp434A0
170G1	16	304	BE964134	Hs.296246	4.00E-96	1	601657818R1 cDNA, 3' end /clone=IMAGE:3876028
597G5	168	1564	NM_014456	Hs.296251	0	18	programmed cell death 4 (PDCD4), mRNA /cds=(84
184A12	686	1564	U96628	Hs.296251	0	2	nuclear antigen H731-like protein mRNA, compl
479H10	247	540	NM_002072	Hs.296261	1.00E-117	1	guanine nucleotide binding protein (G protein
179H11	48	250	BF315059	Hs.296266	3.00E-56	1	601899090F1 5' end /clone=IMAGE:4128334
182E9	1576	2251	AK023460	Hs.296275	0	2	FLJ13398 fis, clone PLACE1001377, highly
459B11	305	545	BF340402	Hs.296317	1.00E-79	1	602036746F1 cDNA, 5' end /clone=IMAGE:4184602
459B12	349	721	AK001838	Hs.296323	0	1	cDNA FLJ10976 fis, clone PLACE1001399 /cds=UN
179F8	1	756	BF342246	Hs.296333	0	2	602013019F1 5' end /clone=IMAGE:4148741

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

171D1	12	330	AV693913	Hs.296339	1.00E-100	1	AV693913 cDNA, 5' end /clone=GKCDVG04 /clone_
39B9	1	297	AB046771	Hs.296350	1.00E-167	1	for KIAA1551 protein, partial cds /cds=(0
36H12	547	1089	M96995	Hs.296381	0	2	epidermal growth factor receptor-binding pro
459F1	867	1020	NM_014499	Hs.296433	4.00E-76	1	putative purinergic receptor (P2Y10), mRNA /c
584A11	615	1287	NM_006392	Hs.296585	0	4	nucleolar protein (KKE/D repeat) (NOP56), mRN
593F7	209	752	NM_005678	Hs.296948	0	2	SNRPN upstream reading frame (SNURF), transcr
174F7	493	681	BE253125	Hs.297095	2.00E-60	5	601116648F1 cDNA, 5' end /clone=IMAGE:3357178
123H9	132	413	BE965554	Hs.297190	9.00E-88	1	601659486R1 cDNA, 3' end /clone=IMAGE:3896204
123D6	1105	1595	AF113676	Hs.297681	0	1	clone FLB2803 PRO0684 mRNA, complete cds /cds=
71C6	1076	1630	NM_003380	Hs.297753	0	2	vimentin (VIM), mRNA /cds=(122,1522) /gb=NM_0
586G5	1179	1452	NM_001908	Hs.297939	1.00E-142	1	cathepsin B (CTSB), mRNA /cds=(177,1196) /gb=
521E7	1	220	NM_001022	Hs.298262	1.00E-119	4	ribosomal protein S19 (RPS19), mRNA /cds=(22,4
466H7	9	339	AW614181	Hs.298654	1.00E-153	1	hg77d03.x1 cDNA, 3' end /clone=IMAGE:2951621
464A4	675	1232	BC001077	Hs.299214	0	1	clone IMAGE:2822295, mRNA, partial cds /cds=
466F3	49	337	AA132448	Hs.299416	1.00E-141	1	zo20a03.s1 cDNA, 3' end /clone=IMAGE:587404 /
589B10	123	339	AW073707	Hs.299581	1.00E-55	30	xb01h03.x1 cDNA, 3' end /clone=IMAGE:2575061
521H4	3	371	NM_001000	Hs.300141	1.00E-125	4	ribosomal protein L39 (RPL39), mRNA /cds=(37,1
599F12	36	328	AW243795	Hs.300220	2.00E-67	1	xo56f02.x1 cDNA, 3' end /clone=IMAGE:2707995
479A6	173	356	AW262077	Hs.300229	3.00E-64	1	xq61e07.x1 cDNA, 3' end /clone=IMAGE:2755140
111C8	806	1350	NM_018579	Hs.300496	1.00E-147	6	mitochondrial solute carrier (LOC51312), mRN
459D8	1	679	NM_014478	Hs.300684	0	1	calcitonin gene-related peptide-receptor co
522C5	98	1360	NM_001154	Hs.300711	0	10	annexin A5 (ANXA5), mRNA /cds=(192,1154) /gb=
596B7	407	750	NM_003130	Hs.300741	2.00E-83	1	sorcin (SRI), mRNA /cds=(12,608) /gb=NM_00313
191A3	210	440	AA788623	Hs.301104	4.00E-34	9	ah29f09.s1 cDNA, 3' end /clone=1240265 /clone
123E1	15	267	BE963194	Hs.301110	1.00E-60	11	601656811R1 cDNA, 3' end /clone=IMAGE:3865731
116F11	346	650	NM_014029	Hs.301175	2.00E-71	2	HSPC022 protein (HSPC022), mRNA /cds=(18,623)
58D4	489	611	AW863111	Hs.301183	8.00E-50	1	MR3-SN0009-010400-101-f02 cDNA /gb=AW863111
122D8	3644	4034	AB037808	Hs.301434	0	1	mRNA for KIAA1387 protein, partial cds /cds=(0
520F11	276	553	BE886472	Hs.301486	1.00E-111	1	601509688F1 cDNA, 5' end /clone=IMAGE:3911301
512E5	71	687	NM_001011	Hs.301547	0	8	ribosomal protein S7 (RPS7), mRNA /cds=(81,665
463F9	168	689	AV702152	Hs.301570	0	1	AV702152 cDNA, 5' end /clone=ADBBFH05 /clone_
117A12	2239	2395	NM_007167	Hs.301637	5.00E-78	1	zinc finger protein 258 (ZNF258), mRNA /cds=(9
190A6	12942	13156	AF155238	Hs.301698	1.00E-114	1	BAC 180i23 chromosome 8 map 8q24.3 beta-galacto
594F12	1409	1841	NM_005442	Hs.301704	0	1	eomesodermin (Xenopus laevis) homolog (EOMES)
116G12	5477	5571	AB033081	Hs.301721	6.00E-47	1	mRNA for KIAA1255 protein, partial cds /cds=(0
123C4	23	579	BE260041	Hs.301809	1.00E-129	4	601150579F1 cDNA, 5' end /clone=IMAGE:3503419
192E12	1458	1854	NM_007145	Hs.301819	0	1	zinc finger protein 146 (ZNF146), mRNA /cds=(8
590G8	1100	1307	AF132197	Hs.301824	3.00E-57	1	PRO1331 mRNA, complete cds /cds=(422,616) /gb
482E5	1764	2139	NM_001295	Hs.301921	0	1	chemokine (C-C motif) receptor 1 (CCR1), mRNA
583C5	4283	4684	NM_014415	Hs.301956	0	1	zinc finger protein (ZNF-U69274), mRNA /cds=(
173G11	645	839	X58529	Hs.302063	1.00E-104	4	rearranged immunoglobulin mRNA for mu heavy chain enh
597D11	30	369	AL137162	Hs.302114	1.00E-150	5	DNA sequence from clone RP5-843L14 on chromosome 20.
191G9	182	353	AC004079	Hs.302183	9.00E-60	1	PAC clone RP1-167F23 from 7p15 /cds=(0,569) /g
473D2	102	333	BF477640	Hs.302447	1.00E-126	1	7r01c05.x1 cDNA /clone=IMAGE /gb=BF477640 /g
479A9	18	267	BE964028	Hs.302585	7.00E-79	1	601657601R1 cDNA, 3' end /clone=IMAGE:3875617
180A5	894	1325	NM_018295	Hs.302981	0	2	hypothetical protein FLJ11000 (FLJ11000), mR

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

593H6	950	1151	X00437	Hs.303157	1.00E-104	1	mRNA for T-cell specific protein /cds=(37,975) /gb=X0
51G12	274	533	BG054649	Hs.303214	1.00E-138	4	7o45b01.x1 cDNA, 3' end /clone=IMAGE:3576912
189B10	785	1024	NM_002138	Hs.303627	1.00E-133	2	heterogeneous nuclear ribonucleoprotein D (
99B11	1	529	NM_002982	Hs.303649	0	51	small inducible cytokine A2 (monocyte chemota
461E1	397	496	AI472078	Hs.303662	2.00E-28	1	tj85h03.x1 cDNA, 3' end /clone=IMAGE:2148341
103A1	359	687	AF130085	Hs.304177	1.00E-151	1	clone FLB8503 PRO2286 mRNA, complete cds /cds
180B11	52	240	AI824522	Hs.304477	4.00E-57	1	tx71d03.x1 cDNA, 3' end /clone=IMAGE:2275013
519A10	1	104	AI880542	Hs.304620	3.00E-26	1	at80h05.x1 cDNA, 3' end /clone=IMAGE:2378361
479F6	331	582	AA873734	Hs.304886	1.00E-131	1	oh55h07.s1 cDNA, 3' end /clone=IMAGE:1470589
176G3	61	324	AI904802	Hs.304919	2.00E-74	1	IL-BT067-190199-037 cDNA /gb=AI904802 /gi=6
471G6	169	397	AW592876	Hs.304925	1.00E-122	1	hg04d05.x1 cDNA, 3' end /clone=IMAGE:2944617
119D11	3	348	AL049282	Hs.306030	1.00E-179	1	mRNA; cDNA DKFZp564M113 (from clone DKFZp564M1
112F7	2398	3008	U80743	Hs.306094	0	1	CAGH32 mRNA, partial cds /cds=(0,1671) /gb=U80
460C1	243	533	NM_001353	Hs.306098	5.00E-71	1	aldo-keto reductase family 1, member C1 (dihy
126A4	469	543	L08048	Hs.306192	2.00E-28	1	non-histone chromosomal protein (HMG-1) retropseudo
119F3	2113	2237	AL096752	Hs.306327	3.00E-60	1	mRNA; cDNA DKFZp434A012 (from clone DKFZp434A0
467F8	1860	2406	AL390039	Hs.307106	0	1	DNA sequence from clone RP13-383K5 on chromosome Xq22
192B12	1	454	X72475	Hs.307183	0	6	H.sapiens mRNA for rearranged Ig kappa light chain variable
116H11	60	402	AF067519	Hs.307357	1.00E-160	1	PITSLRE protein kinase beta SV1 isoform (CDC2L
472D3	150	478	AW975895	Hs.307486	1.00E-124	1	EST388004 cDNA /gb=AW975895 /gi=8167117 /ug=
458B4	87	354	AW206977	Hs.307542	1.00E-143	1	UI-H-BI1-afs-h-11-0-UI.s1 cDNA, 3' end /clon
463A11	181	397	AI057025	Hs.307879	1.00E-69	1	oy75a12.x1 cDNA, 3' end /clone=IMAGE:1671646
479C6	138	403	BE264564	Hs.308154	1.00E-144	1	601192330F1 cDNA, 5' end /clone=IMAGE:3536383
468G10	118	446	AI361642	Hs.309028	0	1	qy86d04.x1 cDNA, 3' end /clone=IMAGE:2018887
461G12	64	466	AI379735	Hs.309117	7.00E-25	1	tc41c11.x1 cDNA, 3' end /clone=IMAGE:2067188
466H8	15	487	AI380278	Hs.309120	0	1	tf99f08.x1 cDNA, 3' end /clone=IMAGE:2107431
477C8	28	187	AI380449	Hs.309122	7.00E-84	1	tg02f12.x1 cDNA, 3' end /clone=IMAGE:2107631
477C9	47	537	AI380687	Hs.309127	0	1	tg03e04.x1 cDNA, 3' end /clone=IMAGE:2107710
465F4	68	631	AI440337	Hs.309279	0	1	tc88b03.x1 cDNA, 3' end /clone=IMAGE:2073197
465G6	313	404	AI475653	Hs.309347	9.00E-31	1	tc93b04.x1 cDNA, 3' end /clone=IMAGE:2073679
465E7	1	340	AI475827	Hs.309349	1.00E-171	2	tc87a05.x1 cDNA, 3' end /clone=IMAGE:2073104
517G11	62	516	AI707809	Hs.309433	1.00E-115	2	as28g09.x1 cDNA, 3' end /clone=IMAGE:2318560
468D11	290	497	AI523766	Hs.309484	1.00E-103	1	tg94f07.x1 cDNA, 3' end /clone=IMAGE:2116453
186F5	77	418	AI569898	Hs.309629	1.00E-81	1	tr57c12.x1 cDNA, 3' end /clone=IMAGE:2222422
116A12	8	158	AI735206	Hs.310333	2.00E-43	1	at07f03.x1 cDNA, 3' end /clone=IMAGE:2354429
126G12	35	170	AI866194	Hs.310948	1.00E-54	1	wl27a03.x1 cDNA, 3' end /clone=IMAGE:2426092
172G8	86	227	AI926251	Hs.311137	3.00E-44	1	wo41h05.x1 cDNA, 3' end /clone=IMAGE:2457945
477D8	1	115	AI968387	Hs.311448	4.00E-42	2	wu02e08.x1 cDNA, 3' end /clone=IMAGE:2515814
462F10	13	220	AW043857	Hs.311783	1.00E-107	1	wy81g04.x1 cDNA, 3' end /clone=IMAGE:2554998
185A9	46	423	AW130007	Hs.312182	1.00E-130	2	xf26f10.x1 cDNA, 3' end /clone=IMAGE:2619211
515F6	34	181	AW148618	Hs.312412	3.00E-58	2	xe99f02.x1 cDNA, 3' end /clone=IMAGE:2616699
583E12	5945	6393	AL133572	Hs.312840	0	1	mRNA; cDNA DKFZp434I0535 (from clone DKFZp434I
471D5	306	411	AW298430	Hs.313413	1.00E-46	1	UI-H-BW0-ajl-c-09-0-UI.s1 cDNA, 3' end /clon
482F7	1	449	AW440965	Hs.313578	0	1	he06d07.x1 cDNA, 3' end /clone=IMAGE:2918221
473B3	179	463	BG150461	Hs.313610	1.00E-135	1	7k01d08.x1 cDNA, 3' end /clone=IMAGE:3443006
479E9	138	434	AW450835	Hs.313715	1.00E-127	1	UI-H-BI3-alf-f-06-0-UI.s1 cDNA, 3' end /clon
71B9	344	577	AI733018	Hs.313929	1.00E-115	1	oh60h01.x5 cDNA, 3' end /clone=IMAGE:1471441
479B6	217	443	AW629176	Hs.314085	2.00E-70	1	hi52a04.x1 cDNA, 3' end /clone=IMAGE:2975886
191F11	55	123	BE255377	Hs.314898	1.00E-26	1	601115405F1 cDNA, 5' end /clone=IMAGE:3355872

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

522F11	14	204	BE962883	Hs.314941	9.00E-83	3	601656423R1 cDNA, 3' end /clone=IMAGE:3856325
195F12	120	363	BE351010	Hs.315050	2.00E-77	1	ht22g04.x1 cDNA, 3' end /clone=IMAGE:3147510
173A5	429	824	BE410105	Hs.315263	1.00E-133	1	601302278F1 cDNA, 5' end /clone=IMAGE:3637002
481B2	1063	1283	NM_006255	Hs.315366	3.00E-72	1	protein kinase C, eta (PRKCH), mRNA /cds=(166,2
459G1	1428	1700	NM_006850	Hs.315463	1.00E-124	1	suppression of tumorigenicity 16 (melanoma di
113H4	22	359	BE901218	Hs.315633	1.00E-127	2	601676034F1 cDNA, 5' end /clone=IMAGE:3958617
583B7	510	754	BE963666	Hs.316047	2.00E-55	2	601656685R1 cDNA, 3' end /clone=IMAGE:3865820
466E10	488	644	AV729160	Hs.316771	1.00E-54	1	AV729160 cDNA, 5' end /clone=HTCCAB04 /clone_
597A6	50	249	AV710763	Hs.316785	4.00E-31	2	AV710763 cDNA, 5' end /clone=CuAAJH09 /clone_
123C3	41	529	BF183507	Hs.318215	1.00E-158	1	601809991R1 cDNA, 3' end /clone=IMAGE:4040470
193E12	15	2274	NM_006074	Hs.318501	0	7	stimulated trans-acting factor (50 kDa) (STAF
165D8	727	1344	BC002867	Hs.318693	0	1	clone IMAGE:3940519, mRNA, partial cds /cds=
49F8	520	1094	M16942	Hs.318720	0	1	MHC class II HLA-DRw53-associated glycoprotein beta-
172E10	310	944	NM_016018	Hs.318725	0	1	CGI-72 protein (LOC51105), mRNA /cds=(69,1400
585B1	51	296	BF696330	Hs.318782	6.00E-90	4	602125273F1 cDNA, 5' end /clone=IMAGE:4281906
45E12	208.	737	NM_000636	Hs.318885	0	7	superoxide dismutase 2, mitochondrial (SOD2)
460G2	409	663	BG106948	Hs.318893	5.00E-96	1	602291361F1 cDNA, 5' end /clone=IMAGE:4386159
480C1	155	325	BF889206	Hs.319926	4.00E-74	1	RC6-TN0073-041200-013-H02 cDNA /gb=BF889206
178F1	1	387	BG112503	Hs.320972	1.00E-133	3	602282105F1 cDNA, 5' end /clone=IMAGE:4369633
176G4	1092	1339	AL110236	Hs.321022	1.00E-136	1	mRNA; cDNA DKFZp566P1124 (from clone DKFZp566P
461H6	1701	2239	NM_024101	Hs.321130	0	1	hypothetical protein MGC2771 (MGC2771), mRNA
513F2	605	1614	AK001111	Hs.321245	0	2	cDNA FLJ10249 fis, clone HEMBB1000725, highly
525B4	9	251	BE871962	Hs.321262	6.00E-98	15	601448005F1 cDNA, 5' end /clone=IMAGE:3852001
467A4	1974	2223	AK026270	Hs.321454	6.00E-87	1	cDNA: FLJ22617 fis, clone HSI05379, highly sim
589F10	39	276	BF970928	Hs.321477	5.00E-77	1	602270204F1 cDNA, 5' end /clone=IMAGE:4358425
125A7	1102	1584	BC000627	Hs.321677	0	1	Signal transducer and activator of transcript
597H3	2786	2920	AL136542	Hs.322456	4.00E-46	2	mRNA; cDNA DKFZp761D0211 (from clone DKFZp761D
465E2	40	107	BE747224	Hs.322643	7.00E-22	1	601580941F1 cDNA, 5' end /clone=IMAGE:3929386
515A12	1	698	AL050376	Hs.322645	0	2	mRNA; cDNA DKFZp586J101 (from clone DKFZp586J1
589H11	26	265	BG283132	Hs.322653	4.00E-79	6	602406784F1 cDNA, 5' end /clone=IMAGE:4518957
586E5	1939	2162	AK025200	Hs.322680	1.00E-120	3	cDNA: FLJ21547 fis, clone COL06206 /cds=UNKNOW
595A2	1	306	BG311130	Hs.322804	2.00E-70	2	ia55a08.y1 cDNA, 5' end /clone_end=5' /gb=BG3
459H11	742	951	BC002746	Hs.322824	1.00E-111	1	Similar to dodecenoyl-Coenzyme A delta isome
64C3	655	887	NM_020368	Hs.322901	1.00E-112	1	disrupter of silencing 10 (SAS10), mRNA /cds=(
591B8	3626	4574	D80006	Hs.322903	0	3	mRNA for KIAA0184 gene, partial cds /cds=(0,2591) /gb
458C3	5106	5198	NM_003035	Hs.323032	3.00E-43	1	TAL1 (SCL) interrupting locus (SIL), mRNA /cds
526B7	2132	2750	NM_024334	Hs.323193	0	2	hypothetical protein MGC3222 (MGC3222), mRNA
167F4	467	731	NM_014953	Hs.323346	1.00E-136	2	KIAA1008 protein (KIAA1008), mRNA /cds=(93,28
194B8	1913	3596	AB051480	Hs.323463	0	9	mRNA for KIAA1693 protein, partial cds /cds=(0
478H9	75	564	BF700502	Hs.323662	0	1	602128860F1 cDNA, 5' end /clone=IMAGE:4285502

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

119B1	1598	2284	NM_014664	Hs.323712	0	2	KIAA0615 gene product (KIAA0615), mRNA /cds=(
167H2	1410	3683	AB046771	Hs.323822	0	4	mRNA for KIAA1551 protein, partial cds /cds=(0
595C12	1	528	NM_021998	Hs.323950	0	6	zinc finger protein 6 (CMPX1) (ZNF6), mRNA /cd
462F1	1	356	AK026836	Hs.324060	1.00E-176	1	cDNA: FLJ23183 fis, clone LNG11477 /cds=(226,7
122D10	217	424	AK026091	Hs.324187	2.00E-83	1	cDNA: FLJ22438 fis, clone HRC09232, highly sim
525B2	1028	3282	AL136739	Hs.324275	0	2	mRNA; cDNA DKFZp434D2111 (from clone DKFZp434D
459B6	3	482	BF668584	Hs.324342	0	1	602123634F1 cDNA, 5' end /clone=IMAGE:4280408
583D10	232	466	NM_021104	Hs.324406	1.00E-130	2	ribosomal protein L41 (RPL41), mRNA /cds=(83,1
118F8	2262	2819	NM_016824	Hs.324470	0	1	adducin 3 (gamma) (ADD3), transcript variant 1
461A5	46	391	AW968541	Hs.324481	1.00E-111	1	EST380617 cDNA /gb=AW968541 /gi=8158382 /ug=
467F11	927	1189	NM_000817	Hs.324784	1.00E-147	1	glutamate decarboxylase 1 (brain, 67kD) (GAD1
103E12	1686	1771	AK024863	Hs.325093	9.00E-42	1	cDNA: FLJ21210 fis, clone COL00479 /cds=UNKNOW
521E11	4276	4689	AB028990	Hs.325530	0	1	mRNA for KIAA1067 protein, partial cds /cds=(0
480A9	112	333	AA760848	Hs.325874	1.00E-108	1	nz14f06.s1 cDNA, 3' end /clone=IMAGE:1287779
71G8	2619	2868	NM_001964	Hs.326035	1.00E-116	1	early growth response 1 (EGR1), mRNA /cds=(270,
593D6	742	3372	NM_004735	Hs.326159	0	4	leucine rich repeat (in FLII) interacting prot
463G9	42	608	AW975482	Hs.326165	0	1	EST387591 cDNA /gb=AW975482 /gi=8166696 /ug=
526B12	2380	2639	U83857	Hs.326247	1.00E-143	2	Aac11 (aac11) mRNA, complete cds /cds=(77,1663) /gb=
36A1	63	338	AA010282	NA	1.00E-116	1	zi08h07.r1 Soares_fetal_liver_spleen_1NFLS_S1 cDNA
459D10	67	164	AA044450	NA	3.00E-47	1	zk55a02.r1 Soares_pregnant_uterus_NbHPU cDNA clone
469E6	1	216	AA069335	NA	1.00E-104	1	zf74e10.r1 Soares_pineal_gland_N3HPG cDNA clone
463B2	4	205	AA077131	NA	4.00E-88	1	Brain cDNA Library cDNA clone 7B08E10
68H9	17	383	AA101212	NA	0	1	endothelial cell 937223 cDNA clone IMAGE:549605 3'
458F3	120	498	AA115345	NA	0	1	zl09f11.r1 Soares_pregnant_uterus_NbHPU cDNA clone
459E6	36	532	AA122297	NA	0	1	zk97a11.r1 Soares_pregnant_uterus_NbHPU cDNA clone
462C5	1	122	AA136584	NA	2.00E-59	1	fetal retina 937202 cDNA clone IMAGE:565899 3'
594A1	60	412	AA149078	NA	0	1	zl45e09.r1 Soares_pregnant_uterus_NbHPU cDNA clone
515A9	329	449	AA182528	NA	2.00E-46	1	NT2 neuronal precursor 937230 cDNA clone
75H4	7	371	AA187234	NA	1.00E-119	1	endothelial cell 937223 cDNA clone IMAGE:624540 3'
73F10	1	544	AA210786	NA	0	1	cDNA clone IMAGE:682976 5'
525D8	1	119	AA214691	NA	6.00E-60	1	Express cDNA library cDNA 5'
37H4	250	401	AA243144	NA	3.00E-48	1	cDNA clone IMAGE:685113 5'
463B10	145	408	AA250809	NA	1.00E-123	1	cDNA clone IMAGE:684374 5'
464E10	1	303	AA251184	NA	1.00E-119	1	cDNA clone IMAGE:684046 5'
477H8	1	123	AA252909	NA	4.00E-58	3	cDNA clone IMAGE:669292 5'
465C3	1	279	AA258979	NA	1.00E-129	1	cDNA clone IMAGE:687151 5'
588G6	275	529	AA280051	NA	2.00E-94	1	cDNA clone IMAGE:705062 5'
465E9	74	429	AA282774	NA	0	1	cDNA clone IMAGE:713136 5'
459E7	49	466	AA283061	NA	0	1	cDNA clone IMAGE:713078 5'
164B4	41	329	AA284232	NA	1.00E-148	2	zc39c01.T7 Soares_senescent_fibroblasts_NbHSF cDNA
461G8	289	532	AA290921	NA	1.00E-123	1	cDNA clone IMAGE:700335 5'
470G7	29	441	AA290993	NA	0	1	cDNA clone IMAGE:700425 5'
500A12	1	519	AA307854	NA	1.00E-174	1	(HCC) cell line cDNA 5' end similar to
471F4	9	326	AA309188	NA	1.00E-153	1	cDNA
194B6	134	467	AA312681	NA	1.00E-163	1	cDNA 5' end

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

69F3	5	321	AA314369	NA	1.00E-176	1	(HCC) cell line II cDNA 5' end similar
67G10	1	171	AA319163	NA	3.00E-64	2	cDNA 5' end
99A5	1	287	AA322158	NA	1.00E-136	1	cDNA 5' end similar to similar to tropomyosin
171B1	13	310	AA332553	NA	1.00E-135	1	cDNA 5' end
485D11	46	210	AA360634	NA	2.00E-75	1	cDNA 5' end
462G2	1	183	AA377352	NA	4.00E-89	2	cDNA 5' end
523A8	1	407	AA397592	NA	0	1	cDNA clone IMAGE:728546 5'
171G10	1	409	AA401648	NA	0	2	cDNA clone IMAGE:726936 5'
100F5	42	172	AA402069	NA	4.00E-60	1	cDNA clone IMAGE:727161 5'
459H7	48	375	AA412436	NA	1.00E-163	1	cDNA clone IMAGE:731446 5'
102A8	25	120	AA418765	NA	1.00E-46	1	cDNA clone IMAGE:767795 5'
73A3	1	424	AA426506	NA	0	1	cDNA clone IMAGE:768117 5'
72E10	1	442	AA427653	NA	0	11	tumor NbHOT cDNA clone IMAGE:770045 5'
72A1	1	261	AA429783	NA	1.00E-142	1	zw57b01.r1 Soares_total_fetus_Nb2HF8_9w cDNA clone
460D12	126	388	AA431959	NA	1.00E-93	1	cDNA clone IMAGE:782188 3'
460B11	1	437	AA454987	NA	0	1	cDNA clone IMAGE:811916 5'
518A8	1	329	AA457757	NA	1.00E-177	1	fetal retina 937202 cDNA clone IMAGE:838756 5'
460F7	47	490	AA460876	NA	0	1	zx69d04.r1 Soares_total_fetus_Nb2HF8_9w cDNA clone
118H12	1	304	AA476568	NA	1.00E-163	1	zx02f11.r1 Soares_total_fetus_Nb2HF8_9w cDNA clone
40F11	1	533	AA479163	NA	0	1	cDNA clone IMAGE:754246 5' similar to gb:X15606
470F3	76	356	AA482019	NA	1.00E-142	1	cDNA clone IMAGE:746046 3'
466C2	1	354	AA490796	NA	1.00E-148	1	cDNA clone IMAGE:824101 5'
464A9	228	364	AA496483	NA	7.00E-71	1	tumor NbHOT cDNA clone IMAGE:755690 5' similar to
123D11	99	297	AA501725	NA	1.00E-103	1	cDNA clone IMAGE:929806 similar to contains Alu
119G10	128	374	AA501934	NA	1.00E-134	1	cDNA clone IMAGE:956346
166A11	19	140	AA516406	NA	1.00E-48	1	cDNA clone IMAGE:923858 3'
36G1	5	480	AA524720	NA	0	1	cDNA clone IMAGE:937468 3'
109H9	37	286	AA573427	NA	1.00E-130	2	cDNA clone IMAGE:1028913 3'
477B2	8	273	AA579400	NA	1.00E-143	1	cDNA clone IMAGE:915561 similar to contains Alu
178C10	1	354	AA588755	NA	1.00E-177	1	cDNA clone IMAGE:1084243 3'
486G7	35	99	AA613460	NA	6.00E-28	1	cDNA clone IMAGE:1144571 similar to contains
472E9	27	389	AA628833	NA	1.00E-119	1	af37g04.s1 Soares_total_fetus_Nb2HF8_9w cDNA clone
100C3	122	505	AA639796	NA	0	1	cDNA clone IMAGE:1159029 3'
518A7	39	226	AA665359	NA	4.00E-83	1	cDNA clone IMAGE:1205697 similar to
473D9	377	446	AA683244	NA	1.00E-30	1	schizo brain S11 cDNA clone IMAGE:971252 3'
523D7	80	502	AA701667	NA	1.00E-158	1	zi43g09.s1 Soares_fetal_liver_spleen_1NFLS_S1 cDNA
472B1	37	130	AA744774	NA	1.00E-35	1	cDNA clone IMAGE:1283731 3'
98C9	10	254	AA748714	NA	1.00E-111	1	cDNA clone IMAGE:1270595 3'
196D7	3	442	AA806222	NA	0	1	cDNA clone IMAGE:1409989 3'
118A8	10	381	AA806766	NA	0	1	cDNA clone IMAGE:1338727 3'
98B3	56	159	AA826572	NA	7.00E-47	1	cDNA clone IMAGE:1416447 3'
154D9	38	405	AA846378	NA	1.00E-164	1	cDNA clone IMAGE:1394232 3'
459C2	1	491	AA909983	NA	0	2	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:1523142 3'
486A7	1	176	AA916990	NA	1.00E-72	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:1527333 3'
460D2	78	537	AA923567	NA	0	1	cDNA clone IMAGE:1536231 3'
105F4	86	390	AA974839	NA	4.00E-94	1	cDNA clone IMAGE:1567639 3'
461H7	295	383	AA974991	NA	2.00E-30	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:1560953 3'
162B1	398	470	AA976045	NA	9.00E-28	1	cDNA clone IMAGE:1558392 3'
53D8	1	422	AA984245	NA	1.00E-162	1	schizo brain S11 cDNA clone IMAGE:1629672 3'
524A5	3568	4037	AB020681	NA	0	1	mRNA for KIAA0874 protein, partial cds Length = 4440

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

174H3	81	271	AB021288	NA	1.00E-101	1	mRNA for beta 2-microglobulin, complete cds Length = 925
115A2	1920	2309	AB034747	NA	0	4	SIMPLE mRNA for small integral membrane protein of lysosome/late endos
39G7	1578	1920	AB040875	NA	1.00E-135	3	hxCT mRNA for cystine/glutamate exchanger, complete cds Length = 2000
149H2	430	713	AB044971	NA	1.00E-158	1	mRNA for nucleolar phosphoprotein Nopp34, complete cds Length = 1005
458F6	780	1235	AB045118	NA	0	1	FRAT2 mRNA, complete cds Length = 2164
459D12	2694	3564	AB045278	NA	0	2	beta3GnT5 mRNA for beta1,3-N-acetylglucosaminyltransferase 5, complete
103H7	1294	1933	AB049881	NA	1.00E-139	1	similar to Macaca fascicularis brain cDNA, clone:QnpA-18828 Length = 2517
102E11	1142	1772	AB050511	NA	0	1	similar to Macaca fascicularis brain cDNA, clone:QnpA-18828 Length = 2518
460C3	798	930	AB050514	NA	9.00E-54	1	similar to Macaca fascicularis brain cDNA, clone:QnpA-18828 Length = 2519
480A10	4649	5183	AB058677	NA	0	1	mRNA for MEGF11 protein (KIAA1781), complete cds Length = 5702
142G10	2251	2430	AB060884	NA	6.00E-44	1	similar to Macaca fascicularis brain cDNA clone:QtrA-13024, full insert sequence
494G5	1585	1998	AF005213	NA	0	1	ankyrin 1 (ANK1) mRNA, complete cds Length = 2651
154C6	520	826	AF005775	NA	1.00E-150	3	caspase-like apoptosis regulatory protein 2 (clarp) mRNA, alternativel
186B6	772	1248	AF039575	NA	0	1	heterogeneous nuclear ribonucleoprotein D0B mRNA, partial cds
471A4	395	611	AF061944	NA	6.00E-84	1	kinase deficient protein KDP mRNA, partial cds Length = 2653
37G5	277	525	AF067529	NA	1.00E-129	1	PITSLRE protein kinase beta SV18 isoform (CDC2L2) mRNA, partial cds
479D1	1270	1570	AF070635	NA	1.00E-144	1	clone 24818 mRNA sequence Length = 1643
491E2	38	226	AF086214	NA	9.00E-74	1	full length insert cDNA clone ZC64D04 Length = 691
517C2	230	465	AF086431	NA	1.00E-113	1	full length insert cDNA clone ZD79H10 Length = 530
593C6	1	359	AF113210	NA	0	5	MSTP030 mRNA, complete cds Length = 1024
191A8	135	1169	AF113213	NA	0	3	MSTP033 mRNA, complete cds Length = 1281
144E9	799	943	AF116679	NA	9.00E-29	1	PRO2003 mRNA, complete cds Length = 1222
106E3	583	1187	AF116702	NA	0	2	PRO2446 mRNA, complete cds Length = 1356
72F8	878	1205	AF130094	NA	1.00E-175	1	clone FLC0165 mRNA sequence Length = 1548
458G9	730	1463	AF157116	NA	0	1	clone 274512, mRNA sequence Length = 2172
139F11	18	229	AF161430	NA	1.00E-115	1	HSPC312 mRNA, partial cds Length = 360
149H10	406	621	AF161455	NA	3.00E-95	2	HSPC337 mRNA, partial cds Length = 1033
68A9	19	243	AF173954	NA	2.00E-27	1	Cloning vector pGEM-URA3, complete sequence Length = 4350
165B7	65	418	AF202092	NA	0	1	PC3-96 mRNA, complete cds Length = 1068
52H1	361	594	AF212226	NA	1.00E-34	1	RPL24 mRNA, complete cds Length = 1474
162H8	52	404	AF212233	NA	1.00E-179	1	microsomal signal peptidase subunit mRNA, complete cds Length = 794
54E10	680	1316	AF212241	NA	0	3	CDA02 mRNA, complete cds Length = 2179
117D8	2052	2482	AF248648	NA	0	3	RNA-binding protein BRUNOL2 mRNA, complete cds Length = 2615
75E3	326	662	AF249845	NA	0	2	isolate Siddi 10 hypervariable region I, mitochondrial sequence
459G12	791	1267	AF260237	NA	0	1	hairy/enhancer of split 6 (HES6) mRNA, complete cds Length = 1286
177F6	1968	2423	AF267856	NA	0	1	HT033 mRNA, complete cds Length = 2972
115G8	996	1399	AF267863	NA	0	1	DC43 mRNA, complete cds Length = 2493
501H3	426	1152	AF279437	NA	0	107	interleukin 22 (IL22) mRNA, complete cds Length = 1167
174B4	900	1332	AF283771	NA	0	2	clone TCBAP0774 mRNA sequence Length = 1814

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

126C7	454	843	AF332864	NA	1.00E-116	2	similar to Mus Ras association domain family 3 protein (Rassf3) mRNA
105A9	232	624	AF333025	NA	1.00E-140	1	prokineticin 2 precursor (PROK2) mRNA, complete cds Length = 1406
186F1	4543	5058	AF347010	NA	0	3	mitochondrion, complete genome Length = 16570
590B12	4684	5053	AF347013	NA	0	1	mitochondrion, complete genome Length = 16566
517H7	4669	5058	AF347015	NA	0	1	mitochondrion, complete genome Length = 16571
596E9	220	295	AI027844	NA	3.00E-34	1	cDNA clone IMAGE:1671612 3'
599B3	608	609	AI039890	NA	1.00E-45	1	ox97d11.x1 Soares_senescent_fibroblasts_NbHSF cDNA
189H9	22	524	AI041828	NA	0	1	oy34b08.x1 Soares_parathyroid_tumor_NbHPA cDNA clone
471F6	63	526	AI084224	NA	0	1	cDNA clone IMAGE:1671418 3'
142E9	6	372	AI091533	NA	1.00E-179	1	oo23d05.x1 Soares_NSF_F8_9W_OT_PA_P_S1 cDNA clone
72D2	65	529	AI131018	NA	0	6	qb82e07.x1 Soares_fetal_heart_NbHH19W cDNA clone
468F6	9	428	AI223400	NA	0	1	cDNA clone IMAGE:1838447 3' similar to TR:O15383
185H1	94	199	AI267714	NA	5.00E-50	1	SB pool 1 cDNA clone IMAGE:2038526
166A9	1	480	AI275205	NA	0	1	cDNA clone IMAGE:1990616 3'
499F2	4	395	AI281442	NA	0	2	cDNA clone IMAGE:1967452 3'
517H5	155	457	AI298509	NA	1.00E-158	1	cDNA clone IMAGE:1896546 3'
144F7	24	364	AI299573	NA	0	1	cDNA clone IMAGE:1900105 3'
519E9	52	408	AI352690	NA	1.00E-180	1	cDNA clone IMAGE:1946884 3'
466F9	172	440	AI361839	NA	1.00E-109	1	cDNA clone IMAGE:2022012 3'
144C9	118	373	AI362793	NA	7.00E-63	1	cDNA clone IMAGE:2018948 3' similar to gb:M60854
464B11	19	455	AI363001	NA	0	1	cDNA clone IMAGE:2018452 3' similar to contains
127B6	40	257	AI370412	NA	6.00E-96	1	cDNA clone IMAGE:1987587 3'
166C4	58	271	AI371227	NA	1.00E-62	1	cDNA clone IMAGE:1987633 3' similar to
467G7	1	450	AI380016	NA	0	1	cDNA clone IMAGE:2109169 3' similar to
466C5	316	497	AI380390	NA	8.00E-44	1	cDNA clone IMAGE:2107088 3'
466B5	200	477	AI381586	NA	1.00E-126	1	cDNA clone IMAGE:2074796 3'
458G10	347	444	AI384128	NA	2.00E-40	1	cDNA clone IMAGE:2088819 3' similar to contains
467A8	415	522	AI391500	NA	1.00E-41	1	cDNA clone IMAGE:2107686 3'
477D1	14	269	AI392705	NA	1.00E-137	2	cDNA clone IMAGE:2109581 3'
467B11	1	293	AI393970	NA	1.00E-122	1	cDNA clone IMAGE:2107950 3'
522D3	250	526	AI419082	NA	1.00E-127	1	cDNA clone IMAGE:2103029 3'
149A11	25	313	AI440491	NA	1.00E-132	1	cDNA clone IMAGE:2073277 3'
471C1	77	215	AI458739	NA	1.00E-50	1	cDNA clone IMAGE:2149471 3' similar to gb:S85655
116E10	162	503	AI469584	NA	1.00E-171	1	cDNA clone IMAGE:2156522 3'
472C8	1	369	AI498316	NA	0	1	cDNA clone IMAGE:2160886 3' similar to TR:Q62717
468E8	2	451	AI523854	NA	3.00E-92	1	cDNA clone IMAGE:2116683 3'
477B5	23	295	AI524624	NA	2.00E-86	1	cDNA clone IMAGE:2075323 3'
193H3	368	489	AI525644	NA	4.00E-34	1	cDNA 5'
66F1	277	436	AI571519	NA	7.00E-84	2	cDNA clone IMAGE:2225079 3' similar to gb:J03909
171A11	225	429	AI581199	NA	1.00E-101	3	cDNA clone IMAGE:2154787 3' similar to
116F2	337	429	AI597917	NA	4.00E-42	1	cDNA clone IMAGE:2258495 3' similar to contains
461G10	9	398	AI627495	NA	1.00E-179	1	cDNA clone IMAGE:2285386 3'
594D11	206	434	AI628930	NA	1.00E-110	1	cDNA clone IMAGE:2281541 3' similar to
489H9	1	507	AI633798	NA	0	4	cDNA clone IMAGE:2242115 3'
171G7	212	431	AI634972	NA	1.00E-103	1	cDNA clone IMAGE:2284157 3'
165C12	270	581	AI651212	NA	1.00E-175	1	cDNA clone IMAGE:2304186 3'
64B3	1	529	AI678099	NA	0	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:2330166 3'
134H3	186	289	AI684022	NA	1.00E-34	1	cDNA clone IMAGE:2267411 3'

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

110B8	169	496	AI688560	NA	1.00E-132	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:2330535 3'
459F2	160	542	AI697756	NA	0	1	cDNA clone IMAGE:2341330 3'
481F11	21	340	AI700738	NA	1.00E-167	1	cDNA clone IMAGE:2343628 3'
488C5	37	533	AI701165	NA	0	4	cDNA clone IMAGE:2340734 3'
104D9	116	241	AI709236	NA	4.00E-60	1	HPLRB6 cDNA clone IMAGE:2353865 3' similar to
112E1	18	576	AI742850	NA	0	1	wg47a05.x1 Soares_NSF_F8_9W_OT_PA_P_S1 cDNA clone
113H12	5	140	AI748827	NA	1.00E-63	1	HPLRB6 cDNA clone IMAGE:2356401 3'
458B8	150	474	AI760353	NA	0	1	cDNA clone IMAGE:2387703 3'
461H11	334	578	AI762870	NA	1.00E-111	1	cDNA clone IMAGE:2397996 3'
458D10	1	465	AI765153	NA	0	1	cDNA clone IMAGE:2393531 3'
38B5	2	295	AI766963	NA	1.00E-140	1	cDNA clone IMAGE:2400693 3'
471A2	320	394	AI796317	NA	2.00E-31	1	cDNA clone IMAGE:2384100 3'
74D10	15	377	AI802547	NA	1.00E-124	2	cDNA clone IMAGE:2186739 3' similar to TR:O15510
482C9	117	409	AI803065	NA	1.00E-164	1	tj47a07.x1 Soares_NSF_F8_9W_OT_PA_P_S1 cDNA clone
480C5	177	517	AI807278	NA	0	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:2357909 3'
175B12	228	513	AI817153	NA	1.00E-132	1	cDNA clone IMAGE:2413005 3'
66E10	14	268	AI858771	NA	1.00E-119	1	cDNA clone IMAGE:2429769 3'
470H6	65	500	AI880607	NA	0	1	HPLRB6 cDNA clone IMAGE:2355013 3'
181D12	7	512	AI884548	NA	0	1	cDNA clone IMAGE:2437818 3' similar to gb:L06797
468H6	52	528	AI884671	NA	0	1	cDNA clone IMAGE:2431488 3'
597C9	284	383	AI904071	NA	1.00E-48	1	cDNA
467C2	206	351	AI917642	NA	2.00E-59	1	cDNA clone IMAGE:2392330 3'
459D1	25	575	AI948513	NA	0	1	cDNA clone IMAGE:2470532 3'
166E11	152	280	AI954499	NA	4.00E-54	1	cDNA clone IMAGE:2550263 3'
493D7	2032	2171	AJ001235	NA	4.00E-29	1	similar to Papio hamadryas ERV-9 like LTR insertion Length = 2240
116B1	1169	1744	AJ009771	NA	0	1	mRNA for putative RING finger protein, partial Length = 3038
137B9	296	407	AJ271637	NA	4.00E-32	1	similar to Elaeis guineensis microsatellite DNA, clone mEgCIR0219
483E6	4250	4492	AJ278191	NA	1.00E-95	1	similar to Mus musculus mRNA for putative mc7 protein (mc7 gene)
144A8	988	1152	AK001163	NA	1.00E-75	1	cDNA FLJ10301 fis, clone NT2RM2000032 Length = 1298
525C11	49	496	AK001451	NA	0	1	cDNA FLJ10589 fis, clone NT2RP2004389
177D9	707	980	AK004265	NA	7.00E-76	1	similar to Mus 18 days embryo cDNA, RIKEN full-length enriched library,
111E10	777	1121	AK004400	NA	1.00E-112	1	similar to Mus 18 days embryo cDNA, RIKEN full-length enriched library,
458G4	650	1259	AK008020	NA	8.00E-86	1	similar to Mus adult male small intestine cDNA, RIKEN full-length enrich
47G7	31	328	AK009988	NA	1.00E-111	1	similar to Mus adult male tongue cDNA, RIKEN full-length enriched librar
69G7	1801	1987	AK012426	NA	5.00E-68	3	similar to Mus 11 days embryo cDNA, RIKEN full-length enriched library,
62C10	1092	1267	AK013164	NA	6.00E-46	2	similar to Mus 10, 11 days embryo cDNA, RIKEN full-length enriched libra
46D9	3243	3564	AK014408	NA	1.00E-104	1	similar to Mus 12 days embryo embryonic body below diaphragm region
178C11	2069	2326	AK016683	NA	9.00E-83	1	similar to Mus adult male testis cDNA, RIKEN full-length enriched librar
102C12	698	1339	AK018758	NA	0	1	similar to Mus adult male liver cDNA, RIKEN full-length enriched library
585B3	1278	1873	AK021925	NA	0	1	cDNA FLJ11863 fis, clone HEMBA1006926 Length = 2029

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

46F3	1377	2006	AK022057	NA	0	1	cDNA FLJ11995 fis, clone HEMBB1001443, highly similar to Rattus norveg
73E7	344	1112	AK023512	NA	0	9	cDNA FLJ13450 fis, clone PLACE1003027, highly similar to Homo sapiens
465B12	681	1338	AK024202	NA	0	1	cDNA FLJ14140 fis, clone MAMMA1002858, highly similar to Rat cMG1
142D12	254	358	AK024740	NA	9.00E-27	1	cDNA: FLJ21087 fis, clone CAS03323 Length = 826
472F7	1330	1623	AK024764	NA	1.00E-164	1	cDNA: FLJ21111 fis, clone CAS05384, highly similar to AF144700 Homo sa
521A3	26	195	AK024976	NA	2.00E-90	1	cDNA: FLJ21323 fis, clone COL02374 Length = 1348
465D1	2091	2255	AK025769	NA	1.00E-74	1	cDNA: FLJ22116 fis, clone HEP18520 Length = 2271
595E9	16	546	AK026264	NA	0	1	cDNA: FLJ22611 fis, clone HSI04961 Length = 1426
103E1	1353	1866	AK026334	NA	1.00E-126	1	cDNA: FLJ22681 fis, clone HSI10693 Length = 1903
524F3	1635	1742	AK026443	NA	9.00E-51	2	cDNA: FLJ22790 fis, clone KAIA2176, highly similar to HUMPMCA
196H10	938	1286	AK026819	NA	6.00E-82	1	cDNA: FLJ23166 fis, clone LNG09880 Length = 1941
172F7	349	738	AK027258	NA	0	1	cDNA: FLJ23605 fis, clone LNG15982, highly similar to AF113539 Homo sa
187B10	1583	2142	AK027260	NA	1.00E-129	1	cDNA: FLJ23607 fis, clone LNG16050 Length = 2560
190F11	76	636	AL042081	NA	0	1	(synonym: htes3) cDNA clone DKFZp434P171 3'
525A9	1	653	AL042370	NA	0	1	(synonym: htes3) cDNA clone DKFZp434A1821 5'
464G8	59	686	AL042376	NA	0	1	(synonym: htes3) cDNA clone DKFZp434A2421 5'
172B12	380	624	AL047171	NA	1.00E-131	1	(synonym: hute1) cDNA clone DKFZp586F2018 5'
193F3	915	1309	AL049305	NA	1.00E-133	1	mRNA; cDNA DKFZp564A186 (from clone DKFZp564A186) Length = 1669
111H8	102	660	AL049356	NA	1.00E-146	1	mRNA; cDNA DKFZp566E233 (from clone DKFZp566E233) Length = 808
526E6	118	551	AL049932	NA	1.00E-147	2	mRNA; cDNA DKFZp564H2416 (from clone DKFZp564H2416) Length = 1865
37C8	707	996	AL050218	NA	1.00E-156	1	mRNA; cDNA DKFZp586I0923 (from clone DKFZp586I0923) Length = 1282
72A9	1235	1391	AL110164	NA	2.00E-70	1	mRNA; cDNA DKFZp586I0324 (from clone DKFZp586I0324) Length = 1705
107C8	1042	1398	AL117644	NA	0	2	mRNA; cDNA DKFZp434M095 (from clone DKFZp434M095) Length = 1455
62E7	1	475	AL120453	NA	1.00E-117	1	(synonym: hamy2) cDNA clone DKFZp761I208 5'
492A7	77	390	AL121406	NA	1.00E-101	1	(synonym: hmel2) cDNA clone DKFZp762G117 5'
598B1	443	812	AL133879	NA	1.00E-172	1	(synonym: hamy2) cDNA clone DKFZp761J0114 5'
458C10	47	351	AL133913	NA	5.00E-76	1	(synonym: hamy2) cDNA clone DKFZp761M2014 5'
98E7	922	2284	AL136558	NA	0	6	mRNA; cDNA DKFZp761B1514 (from clone DKFZp761B1514) Length = 3453
157F6	3511	3847	AL136797	NA	0	1	mRNA; cDNA DKFZp434N031 (from clone DKFZp434N031); complete cds
68B4	1009	1595	AL136932	NA	0	1	mRNA; cDNA DKFZp586H1322 (from clone DKFZp586H1322); complete cds
458B6	278	955	AL137601	NA	0	1	mRNA; cDNA DKFZp434E0811 (from clone DKFZp434E0811); partial cds
172C9	1866	2423	AL137608	NA	0	1	mRNA; cDNA DKFZp434J1111 (from clone DKFZp434J1111); partial cds
72G1	194	474	AL138429	NA	1.00E-151	1	(synonym: htes3) cDNA clone DKFZp434E0629 3'
463H12	12	356	AL513780	NA	1.00E-124	1	cDNA clone CL0BA003ZF07 5 prime
181B6	43	638	AL520535	NA	0	1	cDNA clone CS0DB006YD20 3 prime
69B6	352	858	AL520892	NA	0	1	cDNA clone CS0DB002YG16 5 prime
182A5	119	617	AL521097	NA	0	1	cDNA clone CS0DB001YA13 3 prime
458E9	3	865	AL528020	NA	0	2	cDNA clone CS0DC028YO09 3 prime

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

485C11	1	431	AL532303	NA	0	1	cDNA clone CS0DM014YJ04 5 prime
196G3	78	698	AL532406	NA	0	1	cDNA clone CS0DM014YL03 5 prime
105H4	154	486	AL533737	NA	1.00E-156	1	cDNA clone CS0DF002YH09 5 prime
594G1	337	756	AL534564	NA	0	1	cDNA clone CS0DF004YI09 5 prime
524A9	403	906	AL540260	NA	0	1	cDNA clone CS0DF032YF03 3 prime
118H5	433	532	AL540399	NA	4.00E-39	1	cDNA clone CS0DE001YM08 5 prime
124C2	270	815	AL543900	NA	0	1	cDNA clone CS0DI005YK13 3 prime
471D3	216	403	AL550229	NA	9.00E-49	1	cDNA clone CS0DI039YD11 5 prime
191F2	324	844	AL554506	NA	0	1	cDNA clone CS0DI083YJ17 5 prime
166F6	64	576	AL556016	NA	0	1	cDNA clone CS0DK010YH04 5 prime
467G9	61	401	AL556919	NA	1.00E-138	1	cDNA clone CS0DK012YI02 5 prime
37D7	149	685	AL559029	NA	0	1	cDNA clone CS0DJ010YJ11 5 prime
590B3	76	287	AL559422	NA	1.00E-111	2	cDNA clone CS0DJ013YN07 5 prime
181H2	168	780	AL559555	NA	0	1	cDNA clone CS0DJ013YP21 5 prime
589E3	28	447	AL561074	NA	0	1	cDNA clone CS0DL001YN01 5 prime
487F9	326	739	AL561892	NA	1.00E-149	1	cDNA clone CS0DB006YL04 3 prime
68F10	12	658	AL562895	NA	0	1	cDNA clone CS0DC021YO20 3 prime
157D7	2	108	AL565736	NA	1.00E-28	1	cDNA clone CS0DF007YC06 3 prime
177B1	231	505	AL567986	NA	1.00E-128	1	cDNA clone CS0DF036YI04 3 prime
512E3	627	815	AL575666	NA	1.00E-94	1	cDNA clone CS0DI069YD02 3 prime
112E10	193	623	AL575755	NA	0	1	cDNA clone CS0DI070YG17 3 prime
70H7	197	757	AL576149	NA	0	1	cDNA clone CS0DI072YK21 3 prime
37F1	275	411	AL577970	NA	1.00E-43	1	cDNA clone CS0DK008YK22 3 prime
65D4	278	828	AL578975	NA	0	1	cDNA clone CS0DK012YN01 3 prime
182G2	70	684	AL579745	NA	0	1	cDNA clone CS0DJ003YG20 5 prime
194F9	450	669	AL582354	NA	3.00E-94	1	cDNA clone CS0DL006YH05 3 prime
184F2	27	501	AL583322	NA	2.00E-37	1	cDNA clone CS0DL012YI10 5 prime
40A3	432	638	AL583391	NA	4.00E-83	1	cDNA clone CS0DL012YA12 3 prime
53G7	6	462	AU117298	NA	0	1	sapiens cDNA clone HEMBA1001091 5'
37G7	218	706	AU118159	NA	0	1	sapiens cDNA clone HEMBA1002998 5'
180F9	174	698	AU120731	NA	0	1	sapiens cDNA clone HEMBB1001298 5'
191F1	298	608	AU135154	NA	1.00E-137	1	sapiens cDNA clone PLACE1001348 5'
466G7	11	125	AU158636	NA	1.00E-53	1	sapiens cDNA clone PLACE4000063 3'
67F9	1	453	AV648670	NA	0	2	cDNA clone GLCBLH08 3'
155D6	97	337	AV650434	NA	1.00E-104	1	cDNA clone GLCCEG06 3'
596H6	1	397	AV651615	NA	0	1	cDNA clone GLCCRF09 3'
99D5	41	232	AV653169	NA	6.00E-78	1	cDNA clone GLCDIB01 3'
331C10	33	365	AV654188	NA	1.00E-103	6	cDNA clone GLCDTC01 3'
121A12	70	188	AV659358	NA	3.00E-47	1	cDNA clone GLCFWC05 3'
460G9	69	476	AV687530	NA	0	1	cDNA clone GKCATH08 5'
470F5	1	174	AV689330	NA	2.00E-50	1	cDNA clone GKCDJE03 5'
109E8	71	471	AV705900	NA	0	1	cDNA clone ADBBFE11 5'
166C9	121	226	AV709955	NA	2.00E-26	1	cDNA clone ADCABF08 5'
117F1	69	582	AV710415	NA	0	1	cDNA clone CuAAND10 5'
523C9	41	536	AV716565	NA	0	6	cDNA clone DCBCAF01 5'
103D7	1	164	AV716644	NA	3.00E-77	2	cDNA clone DCBAUG10 5'
195F11	232	459	AV716791	NA	1.00E-113	2	cDNA clone DCBAZC04 5'
63C4	208	421	AV719659	NA	1.00E-101	1	cDNA clone GLCGRA09 5'
496C4	156	563	AV719938	NA	0	1	cDNA clone GLCFUC08 5'
479A1	120	469	AV720984	NA	1.00E-162	1	cDNA clone HTBBIC02 5'
499D6	70	406	AV721008	NA	1.00E-112	4	cDNA clone HTBBHG03 5'
461C8	182	676	AV723437	NA	0	1	cDNA clone HTBBUE10 5'
585G1	173	552	AV724531	NA	0	1	cDNA clone HTBARD04 5'
113B8	1	149	AV724559	NA	3.00E-40	1	cDNA clone HTBCFB08 5'
111H4	497	498	AV724665	NA	0	1	cDNA clone HTBAYG03 5'
458F5	1	534	AV730135	NA	0	1	cDNA clone HTFAHA06 5'
589F6	21	226	AV735258	NA	6.00E-70	1	cDNA clone cdAAIF03 5'
172C8	209	426	AV738173	NA	9.00E-98	1	cDNA clone CBMAHC04 5'
464G3	43	498	AV743635	NA	0	1	cDNA clone CBLBAC03 5'

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

72D4	43	384	AV745692	NA	1.00E-178	2	cDNA clone NPAACB06 5'
592G12	175	571	AV749844	NA	1.00E-176	1	cDNA clone NPCBVG08 5'
169F6	110	250	AV755117	NA	3.00E-28	1	cDNA clone TPAABA12 5'
99H3	200	513	AV755367	NA	1.00E-131	2	cDNA clone BMFAIB02 5'
595G9	399	549	AV756188	NA	2.00E-31	1	cDNA clone BMFABD08 5'
595A12	8	572	AW002985	NA	0	2	cDNA clone IMAGE:2475831 3'
586B7	184	330	AW004905	NA	8.00E-50	1	cDNA clone IMAGE:2565317 3' similar to
591D6	15	436	AW021037	NA	0	1	Cochlea cDNA clone IMAGE:2483601 5'
188F1	135	476	AW021551	NA	0	1	Cochlea cDNA clone IMAGE:2484414 5'
467E8	73	474	AW027160	NA	1.00E-162	1	Soares_thymus_NHFTh cDNA clone IMAGE:2512983 3' similar to
472G2	11	110	AW064187	NA	9.00E-38	1	CD4 intrathymic T-cell cDNA library cDNA 3'
598F3	43	453	AW071894	NA	0	1	cDNA clone IMAGE:2501169 3'
181C7	10	96	AW131768	NA	8.00E-41	1	cDNA clone IMAGE:2619947 3'
181D1	69	216	AW134512	NA	2.00E-77	1	UI-H-BI1-abv-e-05-0-UI.s1 NCI_CGAP_Sub3 cDNA clone IMAGE:2713065 3'
472B10	339	458	AW136717	NA	4.00E-54	1	UI-H-BI1-adm-a-03-0-UI.s1 NCI_CGAP_Sub3 cDNA clone IMAGE:2717092 3'
166B9	240	408	AW137104	NA	6.00E-88	1	UI-H-BI1-acp-e-02-0-UI.s1 NCI_CGAP_Sub3 cDNA clone IMAGE:2714979 3'
188C1	323	461	AW137149	NA	2.00E-72	1	UI-H-BI1-acq-a-05-0-UI.s1 NCI_CGAP_Sub3 cDNA clone IMAGE:2715152 3'
65B2	106	298	AW148765	NA	7.00E-75	1	cDNA clone IMAGE:2616915 3'
524C3	234	429	AW151854	NA	1.00E-76	2	cDNA clone IMAGE:2623546 3' similar to
479A8	6	327	AW161820	NA	1.00E-151	1	brain 00004 cDNA clone IMAGE:2781653 3'
585E10	7	391	AW166442	NA	0	1	Soares_NHCe_cervix cDNA clone IMAGE:2697403 3'
482C6	9	329	AW188398	NA	1.00E-133	1	cDNA clone IMAGE:2665252 3'
522G11	39	516	AW248322	NA	0	1	cDNA clone IMAGE:2820662 5'
473D5	283	416	AW274156	NA	4.00E-69	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:2814367 3'
71C12	20	530	AW293159	NA	0	2	UI-H-BW0-aii-b-08-0-UI.s1 NCI_CGAP_Sub6 cDNA clone IMAGE:2729414 3'
472H11	205	501	AW293424	NA	1.00E-151	1	UI-H-BI2-ahm-a-12-0-UI.s1 NCI_CGAP_Sub4 cDNA clone IMAGE:2727094 3'
465H11	17	124	AW293426	NA	1.00E-48	1	UI-H-BI2-ahm-b-02-0-UI.s1 NCI_CGAP_Sub4 cDNA clone IMAGE:2727122 3'
461H8	19	452	AW295965	NA	0	1	UI-H-BI2-ahh-f-07-0-UI.s1 NCI_CGAP_Sub4 cDNA clone IMAGE:2726917 3'
464B7	250	551	AW300500	NA	3.00E-95	1	cDNA clone IMAGE:2774602 3'
465C7	1	322	AW338115	NA	0	1	cDNA clone IMAGE:2833029 3'
466H5	10	523	AW341449	NA	0	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:2909026 3' similar to
461D9	12	325	AW379049	NA	1.00E-134	1	HT0230 cDNA
186E8	51	277	AW380881	NA	1.00E-103	1	HT0283 cDNA
180D4	260	348	AW384988	NA	2.00E-30	1	HT0427 cDNA
472C1	13	404	AW390233	NA	1.00E-122	1	ST0181 cDNA
462G12	236	321	AW402007	NA	3.00E-40	1	UI-HF-BK0-aao-g-02-0-UI.r1 NIH_MGC_36 cDNA clone IMAGE:3054530 5'
177H2	18	338	AW405863	NA	9.00E-52	1	UI-HF-BL0-acf-e-06-0-UI.r1 NIH_MGC_37 cDNA clone IMAGE:3059026 5'
140G10	6	308	AW440517	NA	1.00E-152	1	cDNA clone IMAGE:2890615 3'
482A10	1	231	AW440869	NA	1.00E-114	1	cDNA clone IMAGE:2918151 3' similar to contains
40B2	18	353	AW444632	NA	4.00E-45	1	UI-H-BI3-ajw-b-11-0-UI.s1 NCI_CGAP_Sub5 cDNA clone IMAGE:2733260 3'
61C2	21	392	AW444812	NA	0	1	UI-H-BI3-ajy-d-11-0-UI.s1 NCI_CGAP_Sub5 cDNA clone IMAGE:2733380 3'
461H10	151	248	AW449610	NA	8.00E-48	1	UI-H-BI3-aku-g-11-0-UI.s1 NCI_CGAP_Sub5 cDNA clone IMAGE:2735804 3'
479E10	9	425	AW451293	NA	0	1	UI-H-BI3-alh-f-06-0-UI.s1 NCI_CGAP_Sub5 cDNA clone IMAGE:2736899 3'

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

489G6	16	303	AW452023	NA	1.00E-125	1	UI-H-BI3-alm-f-06-0-UI.s1 NCI_CGAP_Sub5 cDNA clone IMAGE:2737306 3'
463H8	99	289	AW452096	NA	1.00E-103	1	UI-H-BI3-alo-d-02-0-UI.s1 NCI_CGAP_Sub5 cDNA clone IMAGE:3068186 3'
459B8	71	535	AW499658	NA	0	1	UI-HF-BR0p-ajj-c-07-0-UI.r1 NIH_MGC_52 cDNA clone IMAGE:3074677 5'
37A2	128	395	AW499828	NA	1.00E-110	1	UI-HF-BN0-ake-c-06-0-UI.r1 NIH_MGC_50 cDNA clone IMAGE:3076619 5'
112E5	88	557	AW499829	NA	0	1	UI-HF-BN0-ake-c-07-0-UI.r1 NIH_MGC_50 cDNA clone IMAGE:3076621 5'
523F5	435	517	AW500534	NA	4.00E-36	1	UI-HF-BN0-akj-d-04-0-UI.r1 NIH_MGC_50 cDNA clone IMAGE:3077406 5'
476E10	152	450	AW501528	NA	1.00E-129	1	UI-HF-BP0p-ajf-c-02-0-UI.r1 NIH_MGC_51 cDNA clone IMAGE:3073923 5'
67D10	36	413	AW504212	NA	0	1	UI-HF-BN0-alp-a-11-0-UI.r1 NIH_MGC_50 cDNA clone IMAGE:3080348 5'
100E10	29	364	AW504293	NA	1.00E-159	1	UI-HF-BN0-alg-b-10-0-UI.r1 NIH_MGC_50 cDNA clone IMAGE:3079267 5'
484D12	35	353	AW510795	NA	1.00E-167	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:2911933 3' similar to
480B2	109	446	AW572538	NA	1.00E-162	1	cDNA clone IMAGE:2832030 3'
465D2	272	464	AW573211	NA	2.00E-49	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:2933767 3' similar to
47G6	125	126	AW614193	NA	1.00E-51	1	cDNA clone IMAGE:2951662 3'
499D7	1	341	AW630825	NA	0	2	cDNA clone IMAGE:2969854 5'
62H5	10	423	AW651682	NA	0	2	cDNA clone IMAGE:2901099 5'
104A7	3	461	AW778854	NA	0	1	cDNA clone IMAGE:3037337 3'
484H1	9	453	AW780057	NA	0	1	cDNA clone IMAGE:3036046 3'
491E8	18	348	AW792856	NA	1.00E-164	2	UM0001 cDNA
65D11	64	648	AW810442	NA	0	3	ST0125 cDNA
596F6	49	623	AW813133	NA	0	1	ST0189 cDNA
518H1	131	386	AW819894	NA	1.00E-133	1	ST0294 cDNA
115A7	1	315	AW836389	NA	1.00E-169	3	LT0030 cDNA
486D9	32	237	AW837717	NA	1.00E-65	1	LT0042 cDNA
477B12	84	253	AW837808	NA	4.00E-67	1	LT0042 cDNA
121A11	253	444	AW842489	NA	1.00E-98	1	CN0032 cDNA
472E6	132	447	AW846856	NA	1.00E-149	1	CT0195 cDNA
164F9	1	462	AW856490	NA	0	1	CT0290 cDNA
103C4	23	366	AW859565	NA	0	1	CT0355 cDNA
129D3	81	295	AW866426	NA	1.00E-108	1	SN0024 cDNA
501F9	88	421	AW873028	NA	1.00E-170	3	cDNA clone IMAGE:3120038 3'
98G4	1	294	AW873326	NA	1.00E-107	1	cDNA clone IMAGE:3009400 3'
72D5	55	648	AW886511	NA	0	1	OT0083 cDNA
460A5	101	294	AW891344	NA	1.00E-102	1	NT0079 cDNA
459E9	196	260	AW945538	NA	8.00E-28	1	EN0024 cDNA
479H5	17	224	AW948395	NA	1.00E-102	1	FN0040 cDNA
165E7	2	599	AW949461	NA	0	1	MAGA cDNA
123G9	104	715	AW954112	NA	0	2	MAGC cDNA
183F3	84	503	AW954476	NA	1.00E-159	1	MAGC cDNA
196C6	8	189	AW954580	NA	5.00E-98	1	MAGC cDNA
515H10	1	512	AW955265	NA	0	1	MAGC cDNA
41E8	16	671	AW957139	NA	1.00E-145	2	MAGD cDNA
66A7	335	503	AW958538	NA	4.00E-85	1	MAGE cDNA
465G8	169	615	AW960484	NA	0	1	MAGF cDNA
519E6	44	290	AW960593	NA	1.00E-134	1	MAGF cDNA
594F4	306	571	AW963171	NA	1.00E-112	1	MAGH cDNA
155B2	30	673	AW964218	NA	0	3	MAGH cDNA
173B5	1	553	AW965078	NA	0	1	MAGI cDNA
176A6	7	312	AW965490	NA	1.00E-136	1	MAGI cDNA
498H9	1	456	AW965987	NA	0	2	MAGI cDNA
517D11	105	484	AW966098	NA	0	2	MAGI cDNA

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

166H7	63	559	AW967388	NA	0	1	MAGJ cDNA
462C8	69	212	AW967948	NA	2.00E-72	1	MAGJ cDNA
189C5	8	566	AW968561	NA	0	1	MAGJ cDNA
459C3	129	587	AW969359	NA	0	2	MAGK cDNA
174C1	155	527	AW969546	NA	1.00E-170	1	MAGK cDNA
191F6	158	543	AW973953	NA	1.00E-152	2	MAGM cDNA
461G9	311	437	AW974749	NA	7.00E-47	1	MAGN cDNA
104D1	182	594	AW993791	NA	0	1	BN0034 cDNA
188F5	734	1292	AY007110	NA	0	4	clone TCCCTA00084 mRNA sequence Length = 1656
48D7	692	1169	AY029066	NA	1.00E-76	4	Humanin (HN1) mRNA, complete cds Length = 1567
55B8	1802	2045	BC000141	NA	3.00E-96	1	Similar to myelocytomatosis oncogene, clone MGC:5183, mRNA
37A8	34	301	BC000374	NA	1.00E-101	1	ribosomal protein L18, clone MGC:8373, mRNA, complete cds
178E5	20	551	BC000408	NA	5.00E-53	1	acetyl-Coenzyme A acetyltransferase 2 (acetoacetyl Coenzyme A thiolase
596G2	27	263	BC000449	NA	3.00E-43	2	Similar to ubiquitin C, clone MGC:8448, mRNA, complete cds
179A3	693	1002	BC000514	NA	1.00E-160	3	ribosomal protein L13a, clone MGC:8547, mRNA, complete cds
158F10	169	522	BC000523	NA	1.00E-157	1	Similar to ribosomal protein S24, clone MGC:8595, mRNA, complete cds
515G5	34	270	BC000530	NA	7.00E-38	1	ribosomal protein L19, clone MGC:8653, mRNA, complete cds
39B6	286	1073	BC000590	NA	0	9	actin related protein 2/3 complex, subunit 2 (34 kD), clone MGC:1416,
169A4	929	1314	BC000672	NA	0	1	guanine nucleotide binding protein (G protein), beta polypeptide 2-lik
166H4	1350	1745	BC000771	NA	1.00E-169	8	Similar to tropomyosin 4, clone MGC:3261, mRNA, complete cds
331F9	482	949	BC000967	NA	0	1	clone IMAGE:3449287, mRNA, partial cds Length = 2156
526C6	633	829	BC001169	NA	1.00E-100	1	Similar to esterase 10, clone MGC:1873, mRNA, complete cds
135G12	1598	1766	BC001303	NA	6.00E-42	1	Similar to splicing factor, arginine/serine-rich 2 (SC-35), clone MGC:
491C6	613	714	BC001385	NA	3.00E-34	1	Similar to leucine rich repeat (in FLII) interacting protein 1, clone
108D10	234	641	BC001399	NA	2.00E-79	1	ferritin, heavy polypeptide 1, clone MGC:1749, mRNA, complete cds
196H5	1387	1899	BC001412	NA	6.00E-55	4	eukaryotic translation elongation factor 1 alpha 1, clone MGC:1332, mR
460F5	973	1350	BC001413	NA	0	1	clone IMAGE:3140866, mRNA Length = 1634
520C5	348	472	BC001632	NA	5.00E-34	1	Similar to NADH dehydrogenase (ubiquinone) flavoprotein 2 (24kD), clon
520D10	1729	2205	BC001637	NA	0	2	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit
524A1	564	922	BC001660	NA	1.00E-94	2	ribonuclease 6 precursor, clone MGC:1360, mRNA, complete cds
121E7	275	381	BC001697	NA	2.00E-26	1	Similar to ribosomal protein S15a, clone MGC:2466, mRNA, complete cds
109D1	2441	2835	BC001798	NA	1.00E-123	1	clone MGC:3157, mRNA, complete cds Length = 3041
180D9	741	921	BC001819	NA	5.00E-85	2	ribonuclease 6 precursor, clone MGC:3554, mRNA, complete cds
72H5	1264	2808	BC001854	NA	0	8	methionine adenosyltransferase II, alpha, clone MGC:4537, mRNA, comple
167H8	1099	1436	BC002409	NA	1.00E-49	1	actin, beta, clone MGC:8647, mRNA, complete cds Length = 1858
53H1	2398	2513	BC002538	NA	3.00E-41	1	serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

125B3	246	585	BC002711	NA	1.00E-40	1	cell division cycle 42 (GTP-binding protein, 25kD), clone MGC:3497, mR
331H8	201	557	BC002837	NA	0	1	clone MGC:4175, mRNA, complete cds Length = 1092
150C4	1699	2040	BC002845	NA	8.00E-29	1	eukaryotic translation elongation factor 1 alpha 1, clone MGC:3711, mR
70D7	345	850	BC002900	NA	0	1	Similar to proteasome (prosome, macropain) subunit, alpha type, 2, clo
476B5	1431	1761	BC002929	NA	1.00E-141	1	clone IMAGE:3954899, mRNA, partial cds Length = 2467
38D7	200	688	BC002971	NA	0	2	clone IMAGE:3543711, mRNA, partial cds Length = 1934
74A11	652	1724	BC003063	NA	0	5	Similar to likely ortholog of yeast ARV1, clone IMAGE:3506392, mRNA
105H12	1148	1370	BC003090	NA	1.00E-105	1	COP9 homolog, clone MGC:1297, mRNA, complete cds Length = 1637
50F4	8	301	BC003137	NA	1.00E-115	1	ribosomal protein S3, clone MGC:3657, mRNA, complete cds
175G9	93	216	BC003352	NA	1.00E-33	1	tumor protein, translationally-controlled 1, clone MGC:5308, mRNA, com
587E9	72	554	BC003358	NA	4.00E-60	2	ribosomal protein L10, clone MGC:5189, mRNA, complete cds
71F8	491	911	BC003406	NA	0	1	cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMP-N-acet
512E11	308	372	BC003563	NA	2.00E-27	1	guanine nucleotide binding protein (G protein), gamma 5, clone MGC:196
118B11	76	343	BC003577	NA	1.00E-111	1	clone IMAGE:3544292, mRNA, partial cds Length = 826
107E3	9	634	BC003697	NA	0	1	clone MGC:5564, mRNA, complete cds Length = 2145
128D4	1408	1550	BC004186	NA	1.00E-34	1	guanine nucleotide binding protein, beta 1, clone MGC:2819, mRNA, comp
58H6	554	859	BC004245	NA	1.00E-171	2	ferritin, light polypeptide, clone MGC:10465, mRNA, complete cds
481D8	134	460	BC004258	NA	6.00E-73	1	hypothetical protein PRO1741, clone MGC:10753, mRNA, complete cds
520F6	160	1400	BC004317	NA	0	3	clone MGC:10924, mRNA, complete cds Length = 1837
489G7	511	787	BC004458	NA	2.00E-60	1	enolase 1, (alpha), clone MGC:4315, mRNA, complete cds
115B8	1162	1640	BC004521	NA	0	2	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit
118A2	1126	1369	BC004805	NA	4.00E-38	1	similar to Mus musculus, clone IMAGE:3584831, mRNA Length = 1910
73D2	1174	1751	BC004872	NA	0	1	clone MGC:11034, mRNA, complete cds Length = 2471
522E3	681	993	BC004900	NA	1.00E-175	10	ribosomal protein L13a, clone IMAGE:3545758, mRNA, partial cds
55G12	1	232	BC004928	NA	3.00E-68	1	clone MGC:10493, mRNA, complete cds Length = 2567
520C2	3	139	BC004994	NA	1.00E-31	1	myosin regulatory light chain, clone MGC:4405, mRNA, complete cds
460H4	1577	1923	BC005101	NA	0	1	clone IMAGE:3618561, mRNA Length = 2113
154F12	122	283	BC005128	NA	2.00E-46	1	ribosomal protein L7a, clone MGC:10607, mRNA, complete cds
592C8	647	925	BC005187	NA	2.00E-32	1	Similar to hypothetical protein, clone MGC:12182, mRNA, complete cds
591D1	726	837	BC005361	NA	5.00E-31	1	proteasome (prosome, macropain) subunit, alpha type, 4, clone MGC:1246
458A7	1307	1568	BC005816	NA	4.00E-98	1	Similar to deltex (Drosophila) homolog 1, clone IMAGE:3688330, mRNA, p
122C6	263	378	BC005928	NA	1.00E-29	1	S100 calcium-binding protein A8 (calgranulin A), clone MGC:14536, mRNA
47H11	273	854	BC006008	NA	0	1	clone IMAGE:4285740, mRNA Length = 1040

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

598E1	850	1226	BC006176	NA	0	2	clone IMAGE:4054156, mRNA, partial cds Length = 1423
175A1	570	887	BC006282	NA	1.00E-161	1	Similar to RIKEN cDNA 1110020N13 gene, clone MGC:10540
150H12	543	1098	BC006464	NA	0	1	calmodulin 2 (phosphorylase kinase, delta), clone MGC:2168
583E5	980	1246	BC006849	NA	1.00E-127	1	Similar to RIKEN cDNA 2410044K02 gene, clone MGC:5469
41H7	619	1308	BC007004	NA	0	2	Similar to oxysterol-binding protein-related protein 1, clone IMAGE:40
56C12	13	187	BC007063	NA	6.00E-27	1	peroxiredoxin 1, clone MGC:12514, mRNA, complete cds Length = 973
183C11	2986	3328	BC007203	NA	1.00E-169	1	hypothetical protein MGC10823, clone MGC:12957, mRNA, complete cds
109H10	1343	1627	BC007277	NA	1.00E-156	1	Similar to RIKEN cDNA 0610039P13 gene, clone MGC:15619, mRNA
588E11	423	1324	BC007299	NA	0	3	Similar to ATP synthase, H+ transporting, mitochondrial F1 complex, al
164F12	72	336	BE002854	NA	1.00E-147	1	BN0090 cDNA
106A12	22	608	BE005703	NA	0	1	BN0120 cDNA
472E11	168	297	BE044364	NA	1.00E-66	1	Soares_NFL_T_GBC_S1 cDNA clone IMAGE:3040218 3'
458H11	2	510	BE049439	NA	0	1	cDNA clone IMAGE:2834924 3'
46F7	18	527	BE061115	NA	0	1	BT0041 cDNA
105A8	1	166	BE085539	NA	3.00E-74	1	BT0669 cDNA
467F5	27	247	BE086076	NA	1.00E-115	1	BT0672 cDNA
469B6	5	188	BE091932	NA	6.00E-87	1	BT0733 cDNA
66D7	18	568	BE160822	NA	0	1	HT0422 cDNA
593F8	110	451	BE163106	NA	1.00E-165	1	HT0457 cDNA
468B10	1	461	BE168334	NA	0	1	HT0514 cDNA
192E1	1	602	BE176373	NA	0	1	HT0585 cDNA
109A9	100	377	BE177661	NA	1.00E-129	1	HT0598 cDNA
468B9	27	145	BE178880	NA	3.00E-31	1	HT0609 cDNA
526E11	6	222	BE217848	NA	1.00E-118	3	cDNA clone IMAGE:3174941 3'
115H2	226	227	BE218938	NA	2.00E-97	1	cDNA clone IMAGE:3176478 3'
126B3	1	509	BE222301	NA	1.00E-151	1	cDNA clone IMAGE:3166180 3'
195F2	123	470	BE222392	NA	4.00E-91	1	cDNA clone IMAGE:3166335 3'
170F7	1	375	BE242649	NA	0	1	acute myelogenous leukemia cell (FAB M1) Baylor-HGSC
459F10	35	432	BE247056	NA	5.00E-84	1	cell acute lymphoblastic leukemia Baylor-HGSC project=TCBA
491G11	269	516	BE253336	NA	1.00E-116	1	cDNA clone IMAGE:3357826 5'
471H10	140	202	BE254064	NA	2.00E-26	1	cDNA clone IMAGE:3354554 5'
521H9	22	605	BE292793	NA	0	2	cDNA clone IMAGE:2987838 5'
472A9	33	436	BE297329	NA	0	1	cDNA clone IMAGE:3532809 5'
99E10	59	423	BE328818	NA	0	1	cDNA clone IMAGE:3181355 3'
192C3	4	335	BE348809	NA	0	1	cDNA clone IMAGE:3152438 3'
140G6	206	405	BE348955	NA	3.00E-85	1	cDNA clone IMAGE:3144625 3'
483D12	1	534	BE349148	NA	1.00E-160	1	cDNA clone IMAGE:3150275 3'
491H12	1	526	BE379820	NA	0	1	cDNA clone IMAGE:3510960 5'
481D5	212	333	BE464239	NA	3.00E-45	1	cDNA clone IMAGE:3194693 3'
469H8	31	179	BE466500	NA	2.00E-71	1	cDNA clone IMAGE:3195395 3'
56D11	72	353	BE467470	NA	1.00E-113	1	cDNA clone IMAGE:3212950 3'
471D10	1	249	BE502246	NA	1.00E-119	2	cDNA clone IMAGE:3197344 3'
471C2	255	486	BE502992	NA	1.00E-128	1	cDNA clone IMAGE:3214462 3'
56A2	291	669	BE538333	NA	1.00E-164	1	cDNA clone IMAGE:3454710 5'
191F12	488	587	BE547584	NA	9.00E-28	1	cDNA clone IMAGE:3461312 5'
525F3	5	236	BE550944	NA	1.00E-125	1	cDNA clone IMAGE:3233200 3'
473B7	46	228	BE551867	NA	4.00E-86	1	cDNA clone IMAGE:3195555 3'
467C6	48	404	BE569141	NA	1.00E-162	1	cDNA clone IMAGE:3681180 5'
110D3	193	473	BE613237	NA	1.00E-157	2	cDNA clone IMAGE:3856357 3'

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

140F9	20	344	BE614297	NA	1.00E-84	1	cDNA clone IMAGE:3906037 3'
473B12	63	216	BE645630	NA	3.00E-51	1	cDNA clone IMAGE:3288143 3' similar to contains
460C2	156	594	BE646470	NA	0	1	cDNA clone IMAGE:3292133 3'
172E5	329	491	BE670804	NA	7.00E-72	8	cDNA clone IMAGE:3285031 3' similar to gb:J04130
469D4	50	553	BE674685	NA	0	1	cDNA clone IMAGE:3292800 3' similar to TR:O60688
171F2	10	280	BE676054	NA	1.00E-96	1	cDNA clone IMAGE:3295273 3'
102E12	102	357	BE737348	NA	2.00E-93	1	cDNA clone IMAGE:3640772 5'
121C11	198	488	BE748663	NA	1.00E-150	1	cDNA clone IMAGE:3838675 3'
126D1	208	449	BE763412	NA	1.00E-122	1	NT0036 cDNA
172H5	52	581	BE768647	NA	0	1	FT0010 cDNA
176F12	178	646	BE792125	NA	0	1	cDNA clone IMAGE:3936215 5'
71A6	16	437	BE825187	NA	0	1	CN0028 cDNA
115F11	14	132	BE858152	NA	4.00E-60	1	cDNA clone IMAGE:3306735 3'
61A11	1	448	BE872245	NA	0	1	cDNA clone IMAGE:3850435 5'
171B8	155	377	BE875145	NA	8.00E-88	1	cDNA clone IMAGE:3891244 5'
108A6	370	539	BE876375	NA	7.00E-72	2	cDNA clone IMAGE:3889033 5'
166B1	1	472	BE877115	NA	1.00E-153	1	cDNA clone IMAGE:3887598 5'
63D11	208	496	BE878973	NA	1.00E-141	1	cDNA clone IMAGE:3895002 5'
525C3	208	400	BE879482	NA	7.00E-88	1	cDNA clone IMAGE:3894277 5'
526F7	335	603	BE881113	NA	1.00E-126	1	cDNA clone IMAGE:3894306 5'
152G12	122	659	BE881351	NA	0	2	cDNA clone IMAGE:3892808 5'
589H4	118	510	BE882335	NA	0	2	cDNA clone IMAGE:3907044 5'
51B12	199	631	BE884898	NA	3.00E-56	1	cDNA clone IMAGE:3908551 5'
114C1	286	530	BE887646	NA	1.00E-121	1	cDNA clone IMAGE:3913468 5'
120H2	282	706	BE888744	NA	0	1	cDNA clone IMAGE:3915133 5'
107D11	172	497	BE891242	NA	0	1	cDNA clone IMAGE:3917201 5'
513G4	263	662	BE891269	NA	0	1	cDNA clone IMAGE:3917064 5'
166B8	7	453	BE891928	NA	0	1	cDNA clone IMAGE:3920185 5'
185G9	23	390	BE894437	NA	1.00E-145	1	cDNA clone IMAGE:3918224 5'
189A8	211	485	BE896691	NA	1.00E-82	1	cDNA clone IMAGE:3925062 5'
598A7	78	301	BE897669	NA	1.00E-83	1	cDNA clone IMAGE:3923346 5'
191D9	189	575	BE899595	NA	0	3	cDNA clone IMAGE:3952215 5'
331F2	109	287	BF001438	NA	3.00E-96	2	cDNA clone IMAGE:3313517 3'
192C9	57	419	BF033741	NA	0	1	cDNA clone IMAGE:3857635 5'
117H4	73	454	BF056055	NA	0	1	cDNA clone IMAGE:3443950 3' similar to contains
104B10	6	412	BF058599	NA	1.00E-177	1	cDNA clone IMAGE:3477311 3'
331A12	13	164	BF059133	NA	1.00E-72	1	cDNA clone IMAGE:3480249 3'
40H1	81	507	BF060725	NA	0	1	7j59h07.x1 Soares_NSF_F8_9W_OT_PA_P_S1 cDNA clone
464F1	1	510	BF061421	NA	0	1	7j52c11.x1 Soares_NSF_F8_9W_OT_PA_P_S1 cDNA clone
71E11	1	441	BF105172	NA	0	1	cDNA clone IMAGE:4042560 5'
129D7	92	561	BF116224	NA	0	2	cDNA clone IMAGE:3570793 3'
145E10	83	624	BF131060	NA	0	1	cDNA clone IMAGE:4051731 5'
113B6	105	410	BF194880	NA	1.00E-157	1	cDNA clone IMAGE:3643600 3'
157E9	102	308	BF197153	NA	1.00E-108	2	cDNA clone IMAGE:3561933 3'
127H8	1	173	BF197762	NA	3.00E-92	1	cDNA clone IMAGE:3653139 3'
462D1	29	177	BF221780	NA	7.00E-78	1	cDNA clone IMAGE:3578603 3'
472B8	7	229	BF306204	NA	9.00E-70	1	cDNA clone IMAGE:4138980 5'
62A3	187	612	BF309911	NA	1.00E-162	1	cDNA clone IMAGE:4138171 5'
476G4	316	487	BF330908	NA	5.00E-66	1	BT0333 cDNA
524D1	86	258	BF339088	NA	8.00E-88	1	cDNA clone IMAGE:4182956 5'
58G4	13	606	BF341359	NA	0	2	cDNA clone IMAGE:4149195 5'
480E7	68	288	BF357523	NA	4.00E-97	1	HT0945 cDNA
116C9	8	170	BF364413	NA	2.00E-81	1	NN1068 cDNA
168F4	11	595	BF369763	NA	0	1	GN0120 cDNA
495F1	1	318	BF373638	NA	1.00E-108	2	FT0176 cDNA
98E1	81	499	BF377518	NA	0	2	TN0115 cDNA

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

169C5	17	500	BF380732	NA	0	1	UT0073 cDNA
464E11	12	272	BF432643	NA	1.00E-129	1	cDNA clone IMAGE:3406531 3'
183G2	119	548	BF433058	NA	1.00E-112	1	cDNA clone IMAGE:3565500 3'
473F9	21	411	BF433353	NA	0	1	cDNA clone IMAGE:3703678 3'
117C9	179	462	BF433657	NA	2.00E-99	1	cDNA clone IMAGE:3702965 3' similar to contains
514A3	170	245	BF435621	NA	2.00E-34	2	Lupski_sciatic_nerve cDNA clone IMAGE:3394901 3' similar to
459G8	78	417	BF445405	NA	1.00E-179	1	cDNA clone IMAGE:3699337 3'
483D10	12	474	BF447885	NA	0	1	cDNA clone IMAGE:3706147 3'
519H12	319	394	BF449068	NA	3.00E-27	1	cDNA clone IMAGE:3579069 3'
584H11	78	487	BF475501	NA	7.00E-50	1	Lupski_sciatic_nerve cDNA clone IMAGE:3396242 3'
471G8	214	400	BF478238	NA	9.00E-61	1	cDNA clone IMAGE:3700476 3' similar to contains
109F10	20	329	BF507849	NA	1.00E-172	1	UI-H-BI4-apv-h-02-0-UI.s1 NCI_CGAP_Sub8 cDNA clone IMAGE:3088755 3'
173E10	147	231	BF510393	NA	1.00E-39	1	UI-H-BI4-aon-h-07-0-UI.s1 NCI_CGAP_Sub8 cDNA clone IMAGE:3085669 3'
464D1	32	460	BF513602	NA	1.00E-106	1	UI-H-BW1-amt-a-11-0-UI.s1 NCI_CGAP_Sub7 cDNA clone IMAGE:3070773 3'
118D9	106	248	BF514341	NA	4.00E-46	1	UI-H-BW1-and-h-10-0-UI.s1 NCI_CGAP_Sub7 cDNA clone IMAGE:3082218 3'
462E3	29	197	BF515538	NA	1.00E-87	1	UI-H-BW1-anq-b-09-0-UI.s1 NCI_CGAP_Sub7 cDNA clone IMAGE:3083081 3'
459C7	70	661	BF525720	NA	0	1	cDNA clone IMAGE:4212877 5'
462F8	151	684	BF526421	NA	0	1	cDNA clone IMAGE:4213536 5'
174H6	1	367	BF530382	NA	0	1	cDNA clone IMAGE:4214327 5'
477C5	183	689	BF569545	NA	0	1	cDNA clone IMAGE:4310435 5'
46C3	2	626	BF571362	NA	0	1	cDNA clone IMAGE:4252059 5'
465B1	350	508	BF591040	NA	3.00E-39	1	cDNA clone IMAGE:3319177 3'
477G7	6	127	BF592138	NA	2.00E-57	1	cDNA clone IMAGE:3573334 3'
180B2	53	264	BF593930	NA	1.00E-114	1	nab48e03.x1 Soares_NSF_F8_9W_OT_PA_P_S1 cDNA clone
185F12	139	578	BF663116	NA	0	1	cDNA clone IMAGE:4308392 5'
471F9	77	590	BF667621	NA	0	1	cDNA clone IMAGE:4278888 5'
41D10	16	664	BF668050	NA	0	2	cDNA clone IMAGE:4279827 5'
491G6	87	275	BF670567	NA	1.00E-97	1	cDNA clone IMAGE:4290961 5'
112B4	17	303	BF671020	NA	1.00E-120	1	cDNA clone IMAGE:4292143 5'
194H6	6	196	BF678298	NA	1.00E-100	1	cDNA clone IMAGE:4248916 5'
514H9	96	179	BF691178	NA	2.00E-32	1	cDNA clone IMAGE:4332544 5'
99H1	146	327	BF691895	NA	2.00E-69	1	cDNA clone IMAGE:4333460 5'
465E12	29	681	BF725383	NA	0	1	cDNA (Un-normalized, unamplified): BX cDNA clone
69B10	17	96	BF726114	NA	3.00E-37	1	cDNA (Un-normalized, unamplified): BY cDNA clone
151H10	18	366	BF732404	NA	0	1	cDNA clone IMAGE:3434918 3'
124D2	36	378	BF736784	NA	1.00E-179	1	KT0018 cDNA
463H5	30	152	BF740663	NA	3.00E-56	1	HB0031 cDNA
469D2	164	398	BF744387	NA	6.00E-74	1	BT0636 cDNA
72E1	17	128	BF749089	NA	1.00E-44	3	BN0386 cDNA
98C3	9	515	BF758480	NA	0	1	CT0539 cDNA
46E11	26	162	BF773126	NA	5.00E-57	1	IT0048 cDNA
124C8	32	257	BF773393	NA	1.00E-115	1	IT0039 cDNA
166G8	312	549	BF797348	NA	1.00E-108	1	cDNA clone IMAGE:4340490 5'
146D8	222	288	BF805164	NA	5.00E-29	1	CI0173 cDNA
49G4	99	460	BF813798	NA	0	5	CI0084 cDNA
469F8	31	455	BF816700	NA	4.00E-88	1	CI0128 cDNA
98C1	37	375	BF818594	NA	1.00E-163	1	CI0184 cDNA
62C9	166	359	BF821451	NA	3.00E-28	1	RT0038 cDNA
51F8	28	367	BF827734	NA	1.00E-175	1	HN0025 cDNA
56F7	15	429	BF845167	NA	9.00E-84	1	HT1035 cDNA
476D11	1	303	BF869167	NA	1.00E-165	2	ET0119 cDNA

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

476H4	12	262	BF875575	NA	1.00E-131	2	ET0100 cDNA
68D6	242	452	BF877979	NA	3.00E-98	1	ET0109 cDNA
37C10	1	381	BF897042	NA	0	3	MT0179 cDNA
465B3	63	193	BF898285	NA	5.00E-60	1	MT0229 cDNA
331C7	274	485	BF899464	NA	3.00E-83	1	MT0211 cDNA
72D8	50	334	BF904425	NA	1.00E-152	1	MT0245 cDNA
159F6	333	417	BF906114	NA	2.00E-35	1	MT0267 cDNA
108H5	6	409	BF926187	NA	0	1	NT0193 cDNA
71F9	192	286	BF928644	NA	1.00E-43	1	NT0216 cDNA
481D4	27	334	BF938959	NA	1.00E-102	1	cDNA clone IMAGE:3706689 3'
189B11	69	183	BF939014	NA	4.00E-29	1	cDNA clone IMAGE:3706658 3'
115G2	85	399	BF940103	NA	1.00E-177	1	cDNA clone IMAGE:3439383 3'
463B3	304	449	BF940291	NA	8.00E-62	1	cDNA clone IMAGE:3577096 3'
122G1	8	339	BF950968	NA	1.00E-170	1	NN1186 cDNA
470B4	251	320	BF962743	NA	2.00E-28	1	NN0045 cDNA
516D5	39	208	BF962934	NA	5.00E-69	1	NN0045 cDNA
593G10	242	597	BF965068	NA	1.00E-177	2	cDNA clone IMAGE:4356776 5'
101A1	6	356	BF965438	NA	1.00E-132	1	cDNA clone IMAGE:4356453 5'
477F3	25	653	BF965960	NA	0	1	cDNA clone IMAGE:4365102 5'
588E4	67	562	BF966028	NA	1.00E-134	1	cDNA clone IMAGE:4364887 5'
467F10	11	282	BF966049	NA	1.00E-122	1	cDNA clone IMAGE:4364941 5'
59E12	81	355	BF966269	NA	1.00E-144	1	cDNA clone IMAGE:4375212 5'
480E11	416	755	BF968628	NA	8.00E-41	1	cDNA clone IMAGE:4359351 5'
37H8	200	500	BF968963	NA	1.00E-148	1	cDNA clone IMAGE:4358390 5'
98H5	396	397	BF969990	NA	1.00E-133	1	cDNA clone IMAGE:4360614 5'
597C3	15	571	BF971075	NA	0	1	cDNA clone IMAGE:4358911 5'
101F1	188	305	BF971984	NA	6.00E-42	1	cDNA clone IMAGE:4329095 5'
464H5	246	602	BF980139	NA	0	1	cDNA clone IMAGE:4373963 3'
63B6	130	597	BF981080	NA	0	1	cDNA clone IMAGE:4401411 5'
167A3	223	418	BF981263	NA	1.00E-101	1	cDNA clone IMAGE:4400757 5'
512C12	1	494	BF981634	NA	0	1	cDNA clone IMAGE:4397101 5'
187H7	26	433	BF997765	NA	1.00E-180	2	GN0127 cDNA
458E4	54	242	BG006820	NA	3.00E-62	1	GN0227 cDNA
106A7	1	604	BG024761	NA	0	1	cDNA clone IMAGE:4363858 5'
459H6	1	524	BG026279	NA	0	1	cDNA clone IMAGE:4386607 5'
460B9	264	512	BG028577	NA	1.00E-105	1	cDNA clone IMAGE:4387518 5'
49E9	100	537	BG033909	NA	0	1	cDNA clone IMAGE:4402729 5'
54C10	1	582	BG033953	NA	0	2	cDNA clone IMAGE:4402647 5'
182B3	1	489	BG034799	NA	0	1	cDNA clone IMAGE:4413514 5'
166F8	13	586	BG036101	NA	0	1	cDNA clone IMAGE:4414135 5'
104A12	56	240	BG054966	NA	1.00E-100	1	cDNA clone IMAGE:3441756 3'
171H10	4	269	BG056668	NA	3.00E-85	1	cDNA clone IMAGE:4169714 3'
146G11	13	522	BG057282	NA	0	5	cDNA clone IMAGE:4140477 3' similar to contains
472A11	69	358	BG057892	NA	1.00E-145	1	7f76e08.x1 Lupski_dorsal_root_ganglion cDNA clone
513B4	2	418	BG058599	NA	0	1	cDNA clone IMAGE:4141266 3'
134B4	201	519	BG058739	NA	1.00E-75	4	cDNA clone IMAGE:4140551 3'
163E7	83	327	BG110599	NA	1.00E-126	1	cDNA clone IMAGE:4368492 5'
118A7	180	577	BG110835	NA	0	1	cDNA clone IMAGE:4366502 5'
37F12	38	649	BG111212	NA	0	5	cDNA clone IMAGE:4369233 5'
464A10	57	673	BG111773	NA	0	1	cDNA clone IMAGE:4372861 5'
464A7	56	411	BG118529	NA	1.00E-167	1	cDNA clone IMAGE:4443519 5'
458D8	186	715	BG121288	NA	0	1	cDNA clone IMAGE:4450407 5'
166H12	25	339	BG149747	NA	1.00E-177	1	cDNA clone IMAGE:3367325 3'
51H4	4	224	BG149986	NA	1.00E-121	1	cDNA clone IMAGE:3406766 3'
75G3	70	280	BG150273	NA	1.00E-115	4	cDNA clone IMAGE:3442930 3'
500F10	18	677	BG163237	NA	0	3	cDNA clone IMAGE:4446802 5'
519E4	39	575	BG164898	NA	0	3	cDNA clone IMAGE:4453661 5'
119E5	21	276	BG165998	NA	1.00E-120	1	cDNA clone IMAGE:4456017 5'

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

519B8	29	214	BG166279	NA	5.00E-86	1	cDNA clone IMAGE:4455496 5'
103B8	377	499	BG170647	NA	1.00E-45	1	cDNA clone IMAGE:4426826 5'
470F8	184	307	BG180098	NA	4.00E-63	1	cDNA clone IMAGE:4430875 5'
585C4	4	98	BG230563	NA	5.00E-46	1	cDNA clone IMAGE:4143330 3' similar to contains
48G7	2	298	BG231557	NA	1.00E-119	1	cDNA clone IMAGE:4142471 3'
73C4	188	430	BG231805	NA	1.00E-130	1	cDNA clone IMAGE:4142814 3'
148H4	2	525	BG231961	NA	1.00E-133	12	cDNA clone IMAGE:4143104 3'
484B5	364	533	BG235942	NA	5.00E-81	1	cDNA clone IMAGE:4141389 3'
137B5	97	523	BG236015	NA	6.00E-87	1	cDNA clone IMAGE:4141365 3'
489B11	12	294	BG236084	NA	4.00E-75	2	cDNA clone IMAGE:4141856 3' similar to
45H2	1	492	BG249224	NA	1.00E-139	1	cDNA clone IMAGE:4470038 5'
172F1	1	562	BG254117	NA	0	1	cDNA clone IMAGE:4475233 5'
588F3	66	202	BG254292	NA	9.00E-43	1	cDNA clone IMAGE:4477042 5'
583B5	8	183	BG272304	NA	7.00E-45	1	cDNA clone IMAGE:4257371
73A4	119	311	BG282346	NA	3.00E-42	1	cDNA clone IMAGE:4545131 5'
586A2	99	511	BG283706	NA	1.00E-160	1	cDNA clone IMAGE:4519866 5'
152F12	1	676	BG286649	NA	0	5	cDNA clone IMAGE:4499224 5'
479A12	228	601	BG286817	NA	1.00E-142	1	cDNA clone IMAGE:4500259 5'
99B4	1	449	BG288308	NA	0	2	cDNA clone IMAGE:4512706 5'
584G2	54	468	BG288554	NA	0	1	cDNA clone IMAGE:4517068 5'
464E2	244	549	BG289048	NA	1.00E-159	2	cDNA clone IMAGE:4512868 5'
113H1	149	436	BG289347	NA	1.00E-161	1	cDNA clone IMAGE:4516241 5'
39G6	1	503	BG290577	NA	0	1	cDNA clone IMAGE:4517986 5'
48D8	38	440	BG291970	NA	0	1	cDNA clone IMAGE:4517457 5'
60E7	1	398	BG319445	NA	0	4	Keratinocyte Subtraction Library- Downregulated Transcripts Homo
168C2	3	221	BG319498	NA	1.00E-111	2	Keratinocyte Subtraction Library- Downregulated Transcripts Homo
461B12	1	393	BG387694	NA	0	2	cDNA clone IMAGE:4521084 5'
174G11	3	542	BG391695	NA	0	1	cDNA clone IMAGE:4537243 5'
597A4	164	612	BG396292	NA	0	2	cDNA clone IMAGE:4581548 5'
190B10	469	667	BG397564	NA	3.00E-62	2	cDNA clone IMAGE:4564968 5'
593C3	35	461	BG403635	NA	0	1	cDNA clone IMAGE:4526364 5'
57H10	121	495	BG413494	NA	0	1	7j54e06.x1 Soares_NSF_F8_9W_OT_PA_P_S1 cDNA clone
155G11	119	347	BG424974	NA	3.00E-52	1	cDNA clone IMAGE:4591378 5'
45G3	17	332	BG427404	NA	1.00E-159	1	cDNA clone IMAGE:4612518 5'
185C9	16	185	BG432194	NA	3.00E-62	1	cDNA clone IMAGE:4610035 5'
331D4	60	386	BG434865	NA	1.00E-179	1	cDNA clone IMAGE:4605025 5'
464H12	97	295	BG438232	NA	1.00E-105	1	cDNA clone IMAGE:4622433 5'
521F2	280	534	BG468330	NA	1.00E-111	1	cDNA clone IMAGE:4644153 5'
56F6	167	582	BG473228	NA	0	2	cDNA clone IMAGE:4646938 5'
61G3	8	185	BG473813	NA	2.00E-95	1	cDNA clone IMAGE:4647416 5'
119E9	7	377	BG482798	NA	1.00E-178	3	cDNA clone IMAGE:4616253 5'
125F8	47	318	BG489375	NA	1.00E-149	1	cDNA clone IMAGE:4636634 5'
73H3	55	154	BG493253	NA	5.00E-49	1	cDNA clone IMAGE:4672787 5'
111H9	79	754	BG497765	NA	0	1	cDNA clone IMAGE:4665582 5'
171A10	74	476	BG501063	NA	0	1	cDNA clone IMAGE:4668643 5'
471G1	65	197	BG501895	NA	1.00E-63	1	cDNA clone IMAGE:4654344 5'
111E1	16	181	BG503693	NA	4.00E-85	2	cDNA clone IMAGE:4657381 5'
121B6	77	553	BG505271	NA	0	2	cDNA clone IMAGE:4664028 5'
599F2	379	484	BG505379	NA	3.00E-45	1	cDNA clone IMAGE:4657121 5'
105C1	208	646	BG505961	NA	0	1	cDNA clone IMAGE:4072795 5'
521E10	23	440	BG506168	NA	0	4	cDNA clone IMAGE:4072226 5'
119A5	188	596	BG506472	NA	1.00E-103	1	cDNA clone IMAGE:4070820 5'
479D7	34	308	BG527060	NA	1.00E-121	1	cDNA clone IMAGE:4685209 5'
71H3	27	542	BG527658	NA	0	1	cDNA clone IMAGE:4685854 5'
186A7	2	336	BG531486	NA	5.00E-96	1	cDNA clone IMAGE:4699409 5'
187H11	186	662	BG532345	NA	0	1	cDNA clone IMAGE:4699954 5'
64G4	166	650	BG532470	NA	0	1	cDNA clone IMAGE:4699923 5'

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

486E6	224	561	BG533994	NA	1.00E-168	5	cDNA clone IMAGE:4663102 5'
116F9	188	392	BG536394	NA	7.00E-67	1	cDNA clone IMAGE:4689645 5'
75C7	1	452	BG536641	NA	0	2	cDNA clone IMAGE:4691078 5'
175D10	3	114	BG537502	NA	2.00E-49	1	cDNA clone IMAGE:4690780 5'
599E1	356	659	BG538731	NA	1.00E-111	1	cDNA clone IMAGE:4691392 5'
191H9	80	631	BG541679	NA	0	1	cDNA clone IMAGE:4695805 5'
466A4	1	408	BG542394	NA	0	1	cDNA clone IMAGE:4696046 5'
67G12	29	698	BG547561	NA	0	3	cDNA clone IMAGE:4703738 5'
467B6	60	234	BG547627	NA	3.00E-93	2	cDNA clone IMAGE:4703608 5'
488F8	2041	2132	D10495	NA	9.00E-31	1	mRNA for protein kinase C delta-type, complete cds Length = 2163
525B6	21	222	D17042	NA	1.00E-100	2	HepG2 partial cDNA, clone hmd3f07m5 Length = 222
471E4	2287	2877	D17391	NA	0	2	mRNA for alpha 4(IV) collagen, C-terminal Length = 3558
134D8	561	694	D28589	NA	2.00E-59	1	mRNA (KIAA00167), partial sequence Length = 792
112D1	1614	2159	D30036	NA	0	1	mRNA for phosphatidylinositol transfer protein (PI-TPalpha), complete
98H4	1	357	F11941	NA	1.00E-180	1	brain cDNA cDNA clone c-33f05
585G7	15	264	F13765	NA	1.00E-136	1	(1992) cDNA clone FII112 3'
47D11	1	296	F35665	NA	1.00E-146	1	cDNA clone sH5-000005-0/E06
465F5	34	225	H03298	NA	1.00E-70	1	cDNA clone IMAGE:151865 5'
481A6	43	362	H51796	NA	1.00E-123	1	spleen 1NFLS cDNA clone IMAGE:194250 5'
100E3	116	205	H56344	NA	1.00E-37	1	spleen 1NFLS cDNA clone IMAGE:203711 5' similar to
464F9	10	398	H57221	NA	5.00E-45	2	spleen 1NFLS cDNA clone IMAGE:204710 5'
66C3	10	77	H78395	NA	8.00E-28	1	liver spleen 1NFLS cDNA clone IMAGE:233597 3'
105D11	63	365	H81660	NA	1.00E-154	1	2NbHM cDNA clone IMAGE:249138 5'
60G10	1	189	H86841	NA	1.00E-100	1	cDNA clone IMAGE:220310 5' similar to SP:S44265
470D6	1	314	H92914	NA	1.00E-146	1	Soares_pineal_gland_N3HPG cDNA clone IMAGE:231988 3'
483E5	839	944	K02885	NA	1.00E-26	1	T-cell receptor active beta-chain V-D-J-beta-1.2-C-beta-1 (TCRB) mRNA,
516F5	1753	2047	L11284	NA	1.00E-131	1	Homosapiens ERK activator kinase (MEK1) mRNA Length = 2222
525E11	105	738	L40557	NA	1.00E-112	1	perforin (PRF1) mRNA, 3' end Length = 818
74F1	661	826	M11124	NA	5.00E-41	1	MHC HLA DQ alpha-chain mRNA from DRw9 cell line Length = 835
121E3	1323	1870	M12824	NA	0	4	T-cell differentiation antigen Leu-2/T8 mRNA, partial cds Length = 197
66H2	713	1190	M17783	NA	0	1	glia-derived nexin (GDN) mRNA, 5' end Length = 1191
41A9	698	883	M32577	NA	4.00E-28	1	MHC HLA-DQ beta mRNA, complete cds Length = 1104
478D10	436	605	M55674	NA	4.00E-33	1	(clone M212) phosphoglycerate mutase 2 (muscle specific isozyme) (PGAM
469B8	5	377	N20190	NA	0	1	2NbHM cDNA clone IMAGE:264340 3'
109E4	21	449	N23307	NA	0	2	2NbHM cDNA clone IMAGE:267836 3'
171D9	80	381	N25486	NA	1.00E-147	1	2NbHM cDNA clone IMAGE:264068 5'
73H12	1	398	N27575	NA	1.00E-144	2	2NbHM cDNA clone IMAGE:264499 5'
490A11	25	475	N31700	NA	0	1	2NbHM cDNA clone IMAGE:267025 5'
599D6	185	483	N34261	NA	1.00E-150	1	2NbHM cDNA clone IMAGE:267967 5'
188F3	112	357	N36787	NA	1.00E-107	1	2NbHM cDNA clone IMAGE:273145 3'
465B10	7	558	N49836	NA	0	1	yz08a11.s1 Soares_multiple_sclerosis_2NbHMSP cDNA
40D4	199	575	N58136	NA	1.00E-153	1	spleen 1NFLS cDNA clone IMAGE:247587 3'
183E2	227	366	N80578	NA	2.00E-53	1	Soares_fetal_lung_NbHL19W cDNA clone IMAGE:300873 3' similar to
139G6	9	269	N94511	NA	1.00E-125	1	zb80g04.s1 Soares_senescent_fibroblasts_NbHSF cDNA
126B8	1	256	N99577	NA	1.00E-137	2	spleen 1NFLS cDNA clone IMAGE:295067 5'

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

118A10	893	5056	NC_001807	NA	0	7	mitochondrion, complete genome Length = 16568
41B2	1	471	NM_000873	NA	0	1	intercellular adhesion molecule 2 (ICAM2), mRNA Length = 1035
62A8	1877	1958	NM_000958	NA	1.00E-37	4	prostaglandin E receptor 4 (subtype EP4) (PTGER4), mRNA
179H10	53	265	NM_000983	NA	1.00E-44	1	ribosomal protein L22 (RPL22), mRNA Length = 602
331D3	71	343	NM_001024	NA	1.00E-144	5	ribosomal protein S21 (RPS21), mRNA Length = 343
41G10	3162	3565	NM_001243	NA	3.00E-47	1	tumor necrosis factor receptor superfamily, member 8 (TNFRSF8), mRNA
591E9	1027	1483	NM_002211	NA	0	2	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29
497C6	4946	5064	NM_002460	NA	9.00E-36	2	interferon regulatory factor 4 (IRF4), mRNA Length = 5065
597D8	1232	1461	NM_005356	NA	2.00E-48	1	lymphocyte-specific protein tyrosine kinase (LCK), mRNA Length = 2032
166G2	50	319	NM_005745	NA	2.00E-90	1	accessory proteins BAP31/BAP29 (DXS1357E), mRNA Length = 1314
468D2	3245	3480	NM_011086	NA	8.00E-63	1	similar to Mus phosphoinositide kinase, fyve-containing (Pikfyve), mRNA
599A4	1335	1630	NM_014644	NA	2.00E-69	1	KIAA0477 gene product (KIAA0477), mRNA Length = 5676
69C2	818	1361	NM_014905	NA	0	3	glutaminase (GLS), mRNA Length = 4606
495C6	622	838	NM_015435	NA	1.00E-104	1	double ring-finger protein, Dorfin (DORFIN), mRNA Length = 1640
463D11	480	632	NM_015995	NA	1.00E-77	1	Kruppel-like factor 13 (KLF13), mRNA Length = 1079
49C10	817	964	NM_019604	NA	3.00E-28	1	class-I MHC-restricted T cell associated molecule (CRTAM), mRNA
188E4	390	643	NM_019997	NA	6.00E-79	1	similar to Mus musculus cDNA sequence AB041581 (AB041581)
103H2	1421	1662	NM_021432	NA	3.00E-66	1	similar to Mus RIKEN cDNA 1110020M21 gene (1110020M21Rik)
465G11	1685	1761	NM_021777	NA	1.00E-34	1	a disintegrin and metalloproteinase domain 28 (ADAM28), transcript var
166D8	1265	1951	NM_022152	NA	0	1	PP1201 protein (PP1201), mRNA Length = 2309
459G6	1	123	NM_024567	NA	2.00E-36	1	hypothetical protein FLJ21616 (FLJ21616), mRNA Length = 1858
461G2	667	1182	NM_025977	NA	1.00E-28	1	similar to Mus RIKEN cDNA 2510048L02 gene (2510048L02Rik)
62A5	759	1200	NM_030780	NA	0	1	folate transporter/carrier (LOC81034), mRNA Length = 2534
52C11	1277	1954	NM_030788	NA	0	1	DC-specific transmembrane protein (LOC81501), mRNA Length = 1974
108A7	910	3014	NM_031419	NA	0	4	molecule possessing ankyrin repeats induced by lipopolysaccharide
74E11	47	464	NM_031435	NA	0	1	hypothetical protein DKFZp564I0422 (DKFZP564I0422), mRNA
56B3	1518	1962	NM_031453	NA	1.00E-176	1	hypothetical protein MGC11034 (MGC11034), mRNA Length = 3301
46F2	118	663	NM_031480	NA	1.00E-105	1	hypothetical protein AD034 (AD034), mRNA Length = 2495
192B3	51	290	R11456	NA	1.00E-105	1	spleen 1NFLS cDNA clone IMAGE:129880 5' similar to
458B9	43	359	R64054	NA	1.00E-159	1	cDNA clone IMAGE:139969 5'
169F11	1	429	R85137	NA	0	1	brain N2b4HB55Y cDNA clone IMAGE:180492 5'
465B5	16	392	R88126	NA	1.00E-164	1	cDNA clone IMAGE:186850 5'
477F8	1	525	T77017	NA	0	1	1NIB cDNA clone IMAGE:23326 5'
39G11	162	455	T80378	NA	1.00E-145	1	1NIB cDNA clone IMAGE:24693 5'
107D7	1	371	T80654	NA	0	1	spleen 1NFLS cDNA clone IMAGE:108950 5'
465A1	6	314	T85880	NA	1.00E-114	1	spleen 1NFLS cDNA clone IMAGE:112441 5'
48D12	2300	2533	U08015	NA	1.00E-128	1	NF-ATc mRNA, complete cds Length = 2743

Table 3A, Candidate nucleotide sequences identified using differential cDNA hybridization analysis

121F1	13	380	U46388	NA	1.00E-150	1	cell line Patu 8988t cDNA clone xs425
127B12	3	330	U52054	NA	0	4	S6 H-8 mRNA expressed in chromosome 6-suppressed melanoma cells
487C2	4054	4187	U52682	NA	2.00E-28	1	lymphocyte specific interferon regulatory factor/interferon regulatory
110B3	1404	2081	U53530	NA	0	1	cytoplasmic dynein 1 heavy chain mRNA, partial cds Length = 2694
466C8	34	175	U75805	NA	3.00E-47	1	cDNA clone f46
148G12	1513	1639	U87954	NA	1.00E-27	1	erbB3 binding protein EBP1 mRNA, complete cds Length = 1648
70A4	564	1381	U94359	NA	0	2	glycogenin-2 like mRNA sequence Length = 4066
158E4	843	945	U97075	NA	1.00E-33	1	FLICE-like inhibitory protein short form mRNA, complete cds
459A1	227	446	W00466	NA	1.00E-60	1	2NbHM cDNA clone IMAGE:291193 5'
459A2	60	350	W00491	NA	1.00E-126	1	2NbHM cDNA clone IMAGE:291255 5' similar to
459B1	76	551	W02600	NA	0	1	spleen 1NFLS cDNA clone IMAGE:296099 5'
166C10	10	415	W16552	NA	0	1	Soares_fetal_lung_NbHL19W cDNA clone IMAGE:301703 5'
471C6	3	383	W19201	NA	1.00E-149	1	Soares_fetal_lung_NbHL19W cDNA clone IMAGE:303118 5' similar to
520A8	75	382	W19487	NA	1.00E-154	1	zb36f09.r1 Soares_parathyroid_tumor_NbHPA cDNA clone
459B7	57	158	W25068	NA	9.00E-50	1	Soares_fetal_lung_NbHL19W cDNA clone IMAGE:308696 5'
188D3	39	283	W26193	NA	2.00E-91	1	randomly primed sublibrary cDNA
75B12	8	386	W27656	NA	1.00E-166	1	randomly primed sublibrary cDNA
163F8	74	330	W47229	NA	1.00E-117	1	zc39c01.r1 Soares_senescent_fibroblasts_NbHSF cDNA
478E6	2	322	W56487	NA	3.00E-51	1	zc59c07.r1 Soares_parathyroid_tumor_NbHPA cDNA clone
73H4	76	297	W72392	NA	1.00E-121	1	Soares_fetal_heart_NbHH19W cDNA clone IMAGE:345661 3'
66D5	1	457	W74397	NA	0	3	Soares_fetal_heart_NbHH19W cDNA clone IMAGE:345236 5'
496D4	85	450	W79598	NA	0	1	Soares_fetal_heart_NbHH19W cDNA clone IMAGE:347020 5'
165D1	108	287	W80882	NA	4.00E-94	1	Soares_fetal_heart_NbHH19W cDNA clone IMAGE:347240 5'
463G1	5	406	W86427	NA	0	1	zh61c11.s1 Soares_fetal_liver_spleen_1NFLS_S1 cDNA
469G11	1276	1621	X06180	NA	0	1	mRNA for CD7 antigen (gp40) Length = 1656
113E11	126	885	X65318	NA	0	1	Cloning vector pGEMEX-2 Length = 3995
482E1	921	1168	X79536	NA	1.00E-102	1	mRNA for hnRNPcore protein A1 Length = 1198
123G8	408	848	XM_002068	NA	8.00E-73	1	glutamate-ammonia ligase (glutamine synthase) (GLUL), mRNA
185E1	508	734	XM_002158	NA	1.00E-27	1	proteasome (prosome, macropain) subunit, alpha type, 5 (PSMA5), mRNA
71A9	1131	1252	XM_002269	NA	4.00E-29	1	ARP3 (actin-related protein 3, yeast) homolog (ACTR3), mRNA
49G7	1	257	XM_003189	NA	1.00E-142	3	similar to eukaryotic translation initiation factor 4A, isoform 2 (H.
128B5	783	980	XM_003304	NA	6.00E-41	1	toll-like receptor 2 (TLR2), mRNA Length = 2600
185G10	853	1057	XM_003507	NA	2.00E-26	1	small inducible cytokine subfamily B (Cys-X-Cys), member 5 (epithelial
41C9	588	1221	XM_003593	NA	0	1	CD38 antigen (p45) (CD38), mRNA Length = 1227
156C4	127	270	XM_004020	NA	6.00E-71	1	ribosomal protein S23 (RPS23), mRNA Length = 488
66E2	1344	1577	XM_004500	NA	1.00E-46	1	CD83 antigen (activated B lymphocytes, immunoglobulin superfamily) (CD
61C6	474	987	XM_004611	NA	2.00E-80	1	Ras homolog enriched in brain 2 (RHEB2), mRNA Length = 987
184A7	971	1361	XM_004720	NA	0	1	hypothetical protein FLJ11000 (FLJ11000), mRNA Length = 1680